DOCUMENT RESUME

ED 475 139 TM 034 826

AUTHOR	Taylor, Grace; Shepard, Lorrie; Kinner, Freya; Rosenthal, Justin
TITLE	A Survey of Teachers' Perspectives on High-Stakes Testing in Colorado: What Gets Taught, What Gets Lost. CSE Technical Report.
INSTITUTION	California Univ., Los Angeles. Center for the Study of Evaluation.; National Center for Research on Evaluation, Standards, and Student Testing, Los Angeles, CA.
SPONS AGENCY	Council for Education Policy, Research and Improvement
REPORT NO	CSE-TR-588
PUB DATE	2002-02-00
NOTE	84p.
CONTRACT	R305B960002
PUB TYPE	Reports - Research (143)
EDRS PRICE	EDRS Price MF01/PC04 Plus Postage.
DESCRIPTORS	Elementary Secondary Education; State Programs; *State Standards; Teacher Surveys; *Teachers; *Test Use; *Testing Programs
IDENTIFIERS	Colorado; *Colorado Student Assessment Program

ABSTRACT

Using a random sample of 1,000 Colorado teachers, this study surveyed the effects of standards, the Colorado Student Assessment Program (CSAP), and school report cards on instruction and test-related practices. Findings show that standards were perceived to have greater impact on improving instruction than did testing. Teachers said they aligned their curriculum, instruction, and lessons to the Colorado standards by adding important content. Attention to the state standards improved the quality of writing instruction and focused instruction in reading, probability, geometry, and mathematics problem-solving explanations. The reported effects of CSAP testing were more mixed. Attention to CSAP improved writing instruction but shifted instruction away from social studies and science, increased the time spent on test format practice, and lowered faculty morale. Nine appendixes provide additional information, including survey responses. (Contains 12 tables and 26 references.) (Author/SLD)

Reproductions supplied by EDRS are the best that can be made
from the original document.

CRESST

National Center for Research on Evaluation, Standards, and Student Testing

ED 475 139

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

K. Hurst

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

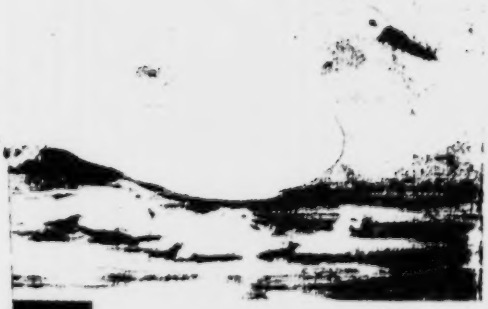
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

A Survey of Teachers' Perspectives on High-Stakes Testing in Colorado: What Gets Taught, What Gets Lost

CSE Technical Report 588

Grace Taylor, Lorrie Shepard, Freya Kinner, and Justin Rosenthal

CRESST/CREDE/University of Colorado at Boulder



TM034826

UCLA Center for the Study of Evaluation

In Collaboration With:

University of Colorado at Boulder • Stanford University • The RAND Corporation
University of Southern California • Educational Testing Service
University of Pittsburgh • University of California, Los Angeles

BEST COPY AVAILABLE

A Survey of Teachers' Perspectives on High-Stakes Testing in Colorado: What Gets Taught, What Gets Lost

CSE Technical Report 588

Grace Taylor, Lorrie Shepard, Freya Kinner,
and Justin Rosenthal
CRESST/CREDE/University of Colorado at Boulder

February 2003

Center for Research on Evaluation, Diversity and Excellence
University of California, Santa Cruz
1156 High Street
Santa Cruz, CA 95064
(408) 459-3500

and

Center for the Study of Evaluation
National Center for Research on Evaluation,
Standards, and Student Testing
Graduate School of Education & Information Studies
University of California, Los Angeles
Los Angeles, CA 90095-1522
(310) 206-1532

Project 2.4 Assessment of Language Minority Students
Lorrie Shepard, Project Director, CRESST/University of Colorado at Boulder

Copyright © 2003 The Regents of the University of California

The work reported herein was supported in part under the Educational Research and Development Centers Program, PR/Award Number R305B960002, as administered by the Office of Educational Research and Improvement, U.S. Department of Education.

The findings and opinions expressed in this report do not reflect the positions or policies of the National Institute on Student Achievement, Curriculum, and Assessment, the Office of Educational Research and Improvement, or the U.S. Department of Education.

A SURVEY OF TEACHERS' PERSPECTIVES ON HIGH-STAKES TESTING IN COLORADO: WHAT GETS TAUGHT, WHAT GETS LOST

Grace Taylor, Lorrie Shepard, Freya Kinner, and Justin Rosenthal
CRESST/CREDE/University of Colorado at Boulder

Abstract

Using a random sample of 1000 Colorado teachers, the researchers in this study surveyed the effects of standards, the Colorado Student Assessment Program (CSAP), and school report cards on instruction and test-related practices. The researchers found that standards were perceived to have a greater impact on improving instruction than did testing. Teachers said they aligned their curriculum, instruction, and lessons to the Colorado standards by adding important content. Specifically, attention to the state standards improved the quality of writing instruction and focused instruction in reading, probability, geometry, and math problem-solving explanations. The reported effects of CSAP testing were more mixed. Attention to CSAP improved writing instruction but shifted instruction away from social studies and science, increased the time spent on test format practice, and lowered faculty morale.

A Survey of Teachers' Perspectives on High-Stakes Testing in Colorado: What Gets Taught, What Gets Lost[1]

Grace Taylor, Lorrie Shepard, Freya Kinner, and Justin Rosenthal

University of Colorado at Boulder
Center for Research on Evaluation, Standards, and Student Testing
Center for Research on Evaluation, Diversity, and Excellence

The development of state assessments and a school accountability system in Colorado closely parallels trends in standards-based reform efforts nationwide. According to the rhetoric of standards-based reform, setting high standards is expected to improve academic achievement by creating higher expectations and thereby focusing greater effort and resources on student learning. However, critics of standards raise a variety of objections, including the fear that higher standards without additional resources may worsen educational inequities or decrease teacher professionalism.

The central role of assessments in standards-based reform has proven to be equally controversial. On the one hand, it is argued that the use of more challenging, open-ended performance assessments, instead of multiple-choice-only, basic-skills tests, will help to better align teaching and learning efforts with ambitious curriculum standards. On the other hand, emphasis on assessments (even good ones) might narrow the curriculum and encourage teachers to teach the test.

Regardless of one's position pro or con in the debate about standards and assessments, it is clear that teachers and teachers' classroom practices are expected to be the key intervening variable that will determine the effects of reforms on student learning. The purpose of the present study was to survey a representative sample of teachers in Colorado to examine the effects of standards, the Colorado Student Assessment Program (CSAP), and school report cards, on instruction and test-related practices.

Standards-Based Reform

[1] The work reported herein was supported under the Educational Research and Development Centers Program, PR/Award Numbers R305B60002 and R306A60001, as administered by the Office of Educational Research and Improvement, U.S. Department of Education.

The findings and opinions expressed in this report do not reflect the positions or policies of the National Institute on Student Achievement, Curriculum, and Assessment, the Office of Educational Research and Improvement, or the U.S. Department of Education.

Standards-based reform began in the United States in the late 1980s. Sometimes referred to as the "second-wave of reform," the standards movement rejected earlier reform agendas, which had focused narrowly on minimum competency testing and mastery of basic skills. The identifying slogan for standard-based reform, "high standards for all students," carries in it a joint commitment to both excellence and equity. Educational reformers like Smith and O'Day (1990) and Resnick and Resnick (1992) argued that the interrelated effects of textbook publishers, who catered to the lowest common denominator to increase sales, and multiple-choice accountability tests had created a de facto national curriculum focused on low-level basic skills. The consequences of this low-level curriculum could be seen in data from the National Assessment of Educational Progress (NAEP) during the 1980s showing a gain in performance on lower-order skills but a decrease in students' problem-solving abilities and conceptual understanding.

To reverse this downward spiral, challenging curriculum standards were needed. Higher standards would refocus instruction more on depth of understanding and the ability of students to reason with and use what they have learned -- rather than on regurgitation of isolated facts. Students should learn to read critically and effectively for multiple purposes, use mathematical knowledge to solve real world problems, write clearly and persuasively using evidence to support their arguments, and understand historical and scientific inquiry. Led first by the National Council of Teachers of Mathematics in 1989, each of the national disciplinary organizations developed curriculum standards to redefine standards of excellence in the discipline and to redirect the goals of instruction. For example, *The National Science Education Standards* developed by the National Research Council (NRC) (1996) emphasized that students should have the opportunity to learn fundamental concepts in depth, to develop subject matter knowledge in the context of inquiry, and to become adept at using scientific knowledge to address societal issues and make personal decisions.

While politicians who called for higher standards were concerned primarily with the economic consequences of poor educational achievement, there is also a substantial body of research on learning showing that students would benefit from higher expectations (Bransford, Brown, and Cocking, 1999). Students learn more and are able to use their knowledge if they are actively trying to figure things out and if they develop their knowledge in a meaningful context of application. Therefore, in some sense, making the curriculum more demanding would enable student learning.

Cognitive research and research on learning likewise supports the commitment to equity in the standards slogan. Scientists no longer believe that the ability to reason and learn is controlled primarily by innate capacity. Rather it is understood that cognitive abilities are "developed" primarily through socially supported interactions. Indeed, the development of intelligence works very much like the development of expertise in academic disciplines. Similarly to processes of acculturation, becoming an expert means learning the ways of thinking and representing problems in a discipline, not just the accumulation of information. The affirmation that "all students can learn" then is a considered, philosophical and research-based commitment intended to counter

BEST COPY AVAILABLE

hereditarian beliefs emphasizing the fixed nature of intellectual abilities. For nearly a century these beliefs had been used to reserve challenging curriculum for only an elite group of high-ability students. This view point, articulated in numerous standards documents, was epitomized by the conclusion of the Malcom Report, *Promises to Keep: Creating High Standards for American Students*, published by the National Education Goals Panel (1993):

> This represents a new way of thinking – a paradigm shift – about American students. The expectation is that students in every school should be able to reach these standards with adequate support and sustained effort. (p. v)

Standards-Based Reform in Colorado

The development of curriculum standards in Colorado was launched by the Colorado Education Reform Act of 1993 also referred to as House Bill 1313. H. B. 1313 called for the development of state-level Model Content Standards in eleven subject areas: reading, writing, mathematics, science, history, geography, civics, art, music, physical education, and foreign language, specifying what students should know and be able to do at three grade levels: grades 4, 8, and 12. School districts were required by January of 1997 to adopt their own content standards in the first priority content areas. District standards were required to be at least as rigorous as the state standards. Districts also were charged with realigning local curriculum and programs of instruction "to ensure that each student will have the educational experiences needed to achieve the adopted content standards."

The preamble of H. B. 1313 stated the legislative intent, which mirrored the national standards agenda:

> Because all children can learn at significantly higher levels than are currently required of them, it is the obligation of the General Assembly, the Department of Education, school districts, educators, and parents to provide children with schools that reflect high expectations and create conditions where these expectations can be met. Through a shared sense of accountability and a cooperative spirit among state government, school districts, educators, parents, business persons, and the community, school districts and educators can develop and teach to high standards which will enable students to achieve world-class knowledge and skills. The General Assembly further declares that this system of standards-based education will serve as an anchor for educational reform, will promote authentic assessment of student learning, will reinforce accountability, and will encourage equity.

Also in keeping with the features of the standards movement nationally, H. B. 1313 mandated the development of a statewide standards-based student assessment program. Intended to be used to monitor progress of the state and school districts in

attainment of the standards, the Colorado Student Assessment Program (CSAP) was developed using a combination of multiple-choice, constructed response, and extended response items aligned with the content standards. Unlike standards rhetoric, however, which called for "capacity building" and the investment of resources to ensure that teachers would be prepared to implement new curricula, Colorado's legislation did not provide for additional resources.

Research on Teaching the Test

If standards-based reform meant rethinking curriculum to engage students in using their knowledge rather than merely repeating what had been "taught," then it was inevitable that assessment would have to change to better reflect the kinds of tasks with which students were engaging. An additional impetus for assessment reform, however, was the evidence of test-score inflation and curriculum distortion gathered from research on teaching the test conducted during the 1980s. Propelled by the dire rhetoric of *A Nation at Risk* (National Commission on Excellence in Education,1983), 38 states had adopted basic-skills testing programs and in many cases mandated high-stakes consequences on students, teachers, or schools when performance was poor. Although initially test-score gains were touted as proof that basic-skills reforms were raising achievement, conflicting evidence from independent tests raised serious questions about whether students were learning more or just getting better at the specific item types represented on the tests.

By the late 1980s, the divergence of trends on NAEP cited previously was one indicator that things might be amiss. In 1987, John Cannell gained considerable public attention by pointing out that all 50 states claimed that their state assessment scores were above the national average. Linn, Graue, and Sanders (1990) conducted a more systematic national study of a representative sample of school districts as well as the 50 states and confirmed that, indeed, nearly all states and a disproportionate number of districts were reporting achievement averages about the national norm. Linn et al. (1990) also compared achievement trends on the six most widely used standardized tests with trends on NAEP and found that scores on the more familiar, and possibly taught-to tests, were increasing at a faster rate than scores on NAEP. These patterns prompted the phrase, "test score inflation."

More in-depth studies of the phenomenon were designed to document teaching practices in high-stakes testing contexts and to evaluate whether the increases in student learning were real or inflated. For example, early in the basic-skills accountability movement, Darling-Hammond and Wise (1985) found that teachers stopped giving essay tests as part of regular instruction so that classroom quizzes would more closely parallel the format of standardized tests given at the end of the year. In a year-long observational study, Smith (1989) found that teachers gave up reading real books, writing, and long-term projects and focused instead on word recognition, recognizing spelling errors, language usage, punctuation, and arithmetic operations. These studies suggested that pressure to do well on tests was narrowing the way the that tested subjects were taught as well as reducing the time spent on non-tested subjects. Teachers surveyed by Shepard

and Dougherty (1991), for example, reported that they had eliminated or greatly reduced time spent on social studies and science to allow more time for reading and math. Yet, ironically, the elimination of social studies and science reduces the likelihood that students will have the chance to use their reading and mathematics skills in real contexts.

Koretz, Linn, Dunbar, and Shepard (1991) went further and actually conducted an experiment to estimate the amount of test score inflation. Koretz et al. first developed alternative tests constructed item-by-item to match district accountability tests but using a slightly more open-ended format.[2] They then administered these independent tests as well as unfamiliar standardized tests to random samples of students in high-stakes accountability settings. Student performance dropped as much as a half standard deviation on the unfamiliar tests suggesting that students did not really know all that they appeared to know on the publicly reported measures.

Assessment Reform

If the aim of the standards movement is to establish more challenging curriculum for all students that is focused on higher-order thinking and depth of understanding, then clearly there must be a fundamental change in how assessments are conceptualized as well. Various terms such as *authentic*, *direct*, and *performance-based* assessment are used in standards parlance to convey the idea that assessment tasks themselves, what students do when they demonstrate their learning, must faithfully reflect important learning goals. The National Academy of Education Panel on Standards-Based Educational Reform recommended, for example, that "*assessments should be compatible with and exemplify the content standards*" (McLaughlin & Shepard, 1995, p. 52, emphasis added). If the curriculum were transformed but assessments continued to include only a narrow portion of what students are intended to learn, then distortions will occur both in the meaning of test results and in subsequent efforts to improve student learning.

A great many standards documents provided sample problems both to illustrate and to enact the reform. For example, The Mathematics Sciences Education Board of the National Research Council developed a set of prototypes for mathematics assessment. Intended for fourth graders, the tasks illustrated how different education would have to be to build the confidence as well as proficiencies needed to do them well. Consistent with the reform's intentions, the tasks call for content knowledge in statistics, geometry, and probability as well as number and operations, call for connections with other academic areas, and promote higher-order thinking but asking students to justify their answers, draw a picture to explain their solution, make predictions, and draw generalizations from their problem solutions. Similarly in science, assessment tasks devised to mirror the standards required students to formulate a question, design and conduct scientific investigations, use tools for data collection, formulate and defend a scientific argument, evaluate alternative explanations on the basis of evidence, and communicate the results of

---

[2] In addition to content matching, the original standardized tests and the alternative independent tests were statistically equated before making comparisons.

a scientific study. In other subject areas as well the content and form of assessments had to be changed to better represent desired learning goals. Especially in writing, assessment of writing abilities might include editing tasks but could no longer be limited to measures of grammar, spelling, and punctuation.

Recognizing the distorting effects of accountability testing in the previous decade some reformers promised that new tests, "tests worth teaching to," would not have the same distorting effects on teaching and learning as had previous low-level tests. After reviewing the research on teaching the test, Resnick and Resnick (1992) articulated the following three principles:

1. You get what you assess.
2. You do not get what you do not assess.
3. Build assessments toward which you want educators to teach. (p. 59)

In defense of the third principle, the Resnicks argued that working to help students perform better on a high-stakes accountability measure is a natural response by teachers. If the assessment activities are identical to good instructional activities then time is not wasted or learning distorted by focus instruction on what gets tested. Others have raised questions, however, (Linn, Baker, & Dunbar, 1991) about whether any test can be so perfect an instantiation of the desired curriculum that teaching the test will be synonymous with good instruction.

Capacity Building

Early in the standards movement attention was paid to "opportunity to learn" standards that would, if met, create the conditions in schools needed for students to reach the standards. Smith and O'Day (1990), who were among some of the earliest advocates for standards-based reform, envisioned a reform that was systemic, affecting all aspects of the educational system, and long-term. They placed considerable emphasis on professional development for both pre-service and in-service teachers and for conditions that would enhance teacher professionalism. The National Council on Education Standards and Testing called for the development of school and system "delivery standards" acknowledging that ambitious goals would not be met without shared responsibility for improvement at both the state and local levels (NCEST, 1992). The Malcom report published in 1993 called for high standards but cautioned that "keeping the promise" would be a long process and would require the commitment of parents, government officials, and community members as well as students, teachers, and school administrators.

The very ambitiousness of the standards movement implied that the standards would not be reached overnight. Although the National Academy of Education Panel on Standards-Based Educational Reform verified that compelling evidence exists to support much higher expectations for students under fundamentally different conditions of teaching and learning, the Panel nonetheless cautioned that the enterprise envisioned was

unprecedented (McLaughlin & Shepard, 1995). Although positive examples exist, the knowledge-base is fragmentary. Considerably more development and on-going evaluation would be needed while attempting to implement these ideas on a larger scale. The Panel emphasized the need to build capacity, especially by providing meaningful professional development for teachers different from one-day in-service sessions:

> For most teachers, accomplishing more significant changes will not happen simply by adding new techniques to their current ideas about teaching and learning. More and different opportunities are needed for teachers and administrators to learn the skills and perspectives fundamental to the success of standards-based reform, which considers knowledge as a dynamic product of teachers and students working together. (McLaughlin & Shepard, 1995, p. 66)

The Panel endorsed the idea of beginning to implement standards-based reforms accompanied by on-going evaluation as the best means to develop a coherent and workable system. However, because many of the attendant problems have not yet been solved, the Panel warned that "More high-stakes applications, with potential serious consequences for students, should be prohibited until major technical problems are resolved and reforms of educational opportunity are more fully in place" (p. 66).

High-Stakes Accountability

At the start of the standards movement, leading researchers and advocates for standards-based reform emphasized the contrast between the new reform, calling for high standards, and previous reform efforts aimed at minimum competencies and basic skills. However, a common thread running without interruption through all three decades of educational reform was the call for accountability. By the end of the 90s, across the nation, the accountability aspect of the reform had become more strident. Politicians were no longer patient with the idea that profound changes in the educational system would require concerted effort over a long period of time. In fact, given turn over in office holders, most governors and legislators entering the new century may not be aware of the original research base or rhetorical claims behind the standards movement. Nor are the new century's political leaders aware of the difference between "world-class," ambitious standards and earlier experience with minimum competency expectations. Therefore, they might be unaware of the difficulty in demanding immediately that everyone meet standards previously met only by privileged and elite groups.

Newer accounts of what standards-based reform is all about emphasize "accountability" as the necessary lever to create incentives for teachers and school personnel to attend to standards and implement new curricula. Researcher Frederick Hess, a proponent of high-stakes accountability, argues that for accountability to have a significant effect on educational quality, "educators must be rewarded or sanctioned on the basis of student performance" (Hess, 2001, p. 4-5). Hess recognizes that such a system runs counter to traditional values of the American educational system, which relied on the good will of teachers and intrinsic motivation of students, but insists that

only a coercive accountability system will be effective in transforming the quality of public schooling.

Recent findings reported in the new Title I report published by the National Research Council (NRC) (Elmore & Rothman, 1999) suggest that external incentive systems alone are not likely to lead to instructional improvements. For example, in one study where failing schools were threatened with reconstitution, school personnel focused on short-term gains in test scores rather than trying to make fundamental changes in instructional practices (O'Day, in press). The NRC Title I Panel concluded that accountability is more likely to have positive effects on student learning when accompanied by systematic local efforts to revise curricular and instructional practices. Despite cautions from the Panel, that professional development and instructional improvements are essential to ensure beneficial outcomes from accountability mandates, the current trend nationally is to focus on testing and school report cards as the primary means to improve public education.

School Accountability in Colorado

Consistent with the national trend toward high-stakes accountability, in 2000 Colorado adopted school-report-card legislation, Senate Bill 186. According to the provisions of S.B. 00-186, the content of the state assessment program, CSAP, was intended to remain the same, but the frequency of testing was substantially increased. In addition, the basic purpose and tenor of the assessment program was shifted. Instead of a program-monitoring assessment aimed only at milestone grade levels, CSAP was expanded to include annual testing of all grade levels from grades 3 to 10 in reading, writing, and mathematics. Moreover, CSAP results would now be used to assign an academic performance grade to each public school. The original legislation called for letter grades from A to F derived from a normal distribution of school average scores with the top 8 percent of schools to receive "A"s and the bottom 2 percent to receive "F"s. Subsequently, because of the controversy associated with letter grades, the reporting categories were changed to descriptive rating labels: Excellent, High, Average, Low, and Unsatisfactory. Schools receiving an "Unsatisfactory" for three consecutive years will be taken over by the state and made into charter schools.

S.B. 00-186 was acknowledged to be Governor Bill Owens's plan. The governor believed that giving low grades to low performing schools would cause the school community to rally. Parents and business leaders would become involved and make sure that school performance improved (Owens, 1999). Critics saw S.B. 00-186 as an empty attempt at reform, because it did not provide resources to struggling schools (Pascoe, 2000). They complained that efforts to discredit the public schools were intended to pave the way for vouchers.

Conceptual Framework and Survey Design

The underlying assumptions guiding standards-based reform were used to frame questions asked of Colorado teachers in both written and telephone surveys. The links –

between issues identified in the research literature and question sets used in the study – are summarized as follows:

Standards-based reform. Teachers were asked to report on changes in instruction that might be expected from standards-based reform and accompanying CSAP testing. For example, they were asked about instructional practices associated with more demanding content standards, such as teaching for higher-order thinking skills, and they were asked whether attention to standards increased effort focused on low performing students as well as on middle and high performing students.

Teaching the test and assessment reform. Teachers were asked to report on instructional practices that have increased or decreased because of CSAP. They were also asked specific questions about the amount and type of test preparation activities. Some questions reflected the older research literature, when teaching the test meant eliminating important content and offering only rote practice on test-like formats. Other questions probed the reformers' claim that more ambitious test content, such as extended writing exercises and explaining solutions to math problems, would drive instruction in a positive direction consistent with content standards.

Capacity building. Teachers were also asked to report on the availability of professional development opportunities and other resources needed to ensure the implementation of standards. Have the reforms in Colorado attended to the needs for capacity building and professional development called for in the research literature? Have sufficient resources been focused on poor students and poor schools to make it likely that Colorado's educational reform will achieve the equity goals originally envisioned by standards rhetoric?

High-stakes accountability. Finally, because accountability systems are based on a theory about motivation and the effects of external incentives, teachers were asked for their opinions about the likely motivational effects of school report cards as well as effects on the allocation of resources.

Methods

In this section, study methods are described, which include sample design, instrument development, data collection, and both quantitative and qualitative methods of data analysis.

Data Collection Schedule

Data collection for both the written and telephone surveys was scheduled to begin after the CSAP test was administered for the 2000-2001 school year. The CSAP testing window for third grade was from February 12 to 23, 2001 for third grade students. Fourth through tenth grade students completed the tests between February 12 and March 9, 2001. Nearly all Colorado third through tenth grade students completed CSAP testing by March 9, 2001 (except some students at schools with year-round schedules). Written surveys were mailed between March 16 and March 20, 2001, and all responses were received between March 20 and June 12, 2001. Telephone surveys were administered between March 19 and May 18, 2001.

Sample Design and Implementation

Samples were drawn using a two-stage stratified cluster design. Complete lists of teachers by school are not available for the state as a whole. Therefore, in Stage 1 of the survey design, districts were selected as clusters to make it possible to obtain teacher lists. A sampling frame for districts was created by crossing three levels of average district CSAP scores with three levels of district size for a total of nine strata. Stage 2 involved a random selection of teachers with probability of selection proportionate to the population size within each cluster. One stratum was removed because there were no Colorado school districts classified with those particular CSAP scores and teacher population size (this was a "large" district with "moderate" CSAP scores). Finally, one additional stratum was added to the sampling framework for districts with no CSAP score information. Each district in the non-CSAP stratum was considered a "small" school district. All contacted districts but one agreed to take part in this study. The district that chose not to participate was not replaced in the sampling design; but teachers in other districts in that stratum were weighted proportionately to represent districts of that type.

Recruitment of districts. Both telephone and written survey design stratification were based on two criteria: district CSAP scores and district size. The Colorado Department of Education supplied a data set including the number of full-time employed teachers by district in Colorado, based on the 1999-2000 school year figures. These data were used to stratify districts by size, based on natural breaks in district sizes. Large districts (with teacher populations greater than or equal to 1,925 teachers) were labeled as size "Large," moderately-sized districts (with teacher populations greater than 670 and less than 1,925 teachers) were labeled as size "Medium," and small districts (with teacher populations smaller than or equal to 670 teachers) were labeled as size "Small" (See Table 1).

Table 1

Description of districts categorized as "Large," "Medium," and "Small."

	Number of Colorado teachers (% of Colorado teachers)	Number of Colorado school districts
"Small" districts	12,089 (26.0%)	156
"Medium" districts	19,069 (45.3%)	17
"Large" districts	10,930 (28.7%)	3
TOTALS	42,088 teachers	176 districts

CSAP scores were aggregated across percent proficient for fourth grade writing and eighth grade mathematics using the 2000 data. Percent proficient or better for writing and percent proficient or better for mathematics were added to create a composite raw score for each school district. The raw scores were then rescaled into composite CSAP z-scores. Using this method, school districts were separated into thirds. "High" districts have z-scores greater than or equal to 0.3024. "Low" districts have z-scores less than or equal to -0.5289. "Moderate" districts include those with z-scores less than 0.3024 and greater than -0.5289. Districts with no CSAP information were included in a separate stratum (See Table 2 for more information).

Table 2

Description of districts categorized as receiving "High," "Moderate," and "Low" CSAP scores.

	Number of Colorado teachers (% of Colorado teachers)	Number of Colorado school districts
"High" CSAP districts	20,031 (47.59%)	50
"Moderate" CSAP districts	10,235 (24.32%)	50
"Low" CSAP districts	11,368 (27.01%)	50
Districts with no CSAP information	454 (1.08%)	26
TOTALS	42,088 teachers	176 districts

Names of teachers teaching at public schools are considered public information. Some districts have this information available on the World Wide Web. Other districts provided this information upon request. Although technically it was not necessary to request permission from districts to contact teachers at their school addresses, we wanted superintendents to be aware of the study being conducted, and many superintendents reported in turn that they would inform principals in their district. Lorrie Shepard gave an overview of the study methods and purpose at a meeting of the Denver Area School Superintendents' Council (DASSC) in February 2001 attended by 14 district superintendents. Superintendents of all other sampled districts were contacted individually by telephone. We did not contact teachers in the one district where the superintendent declined to participate.

The final district sampling framework is shown in Table 3. District response rates were calculated by dividing the number of districts agreeing to take part in our study by the intended number of sampled districts in that stratum.

Table 3

Description of districts categorized as receiving "High," "Moderate," and "Low" CSAP scores.

Strata number	Number of Colorado districts in strata	District size	CSAP score	# of sampled districts in stratum (%)	District response rate per stratum
1	2	Large	High	2 (100%)	100%
2	1	Large	Low	1 (100%)	100%
3	8	Medium	High	5 (62.5%)	100%
4	5	Medium	Moderate	3 (60%)	67%
5	4	Medium	Low	3 (75%)	100%
6	40	Small	High	6 (15%)	100%
7	45	Small	Moderate	6 (13.3%)	100%
8	45	Small	Low	5 (11.1%)	100%
9	26	Small	N/A*	2 (7.7%)	100%

Sampling of teachers within districts. Lists of full-time employed teachers were obtained from each sampled district. Within each cluster (district), teachers were randomly selected with the target size of the sample proportional to population. The teacher sampling framework is as follows (see Table 4):

Table 4

Teacher sampling framework per stratum.

Stratum number	Number of Colorado teachers in stratum	Written-Survey # of sampled teachers in stratum (%)	Phone-Survey # of sampled teachers in stratum (%)
1	6,882	131 (1.9%)	33 (0.5%)
2	4,048	77 (1.9%)	21 (0.5%)
3	9,218	175 (1.9%)	16 (0.2%)
4	5,979	114 (1.9%)	24 (0.4%)
5	3,872	83 (2.1%)	18 (0.5%)
6	3,931	74 (1.9%)	19 (0.5%)
7	4,256	81 (1.9%)	20 (0.5%)
8	3,448	66 (1.9%)	16 (0.5%)
9	454	8 (1.8%)	3 (0.7%)
TOTALS	42,088	809	200

* There were no CSAP data available for districts in Stratum 9.

A total of 809 teachers were sampled for the written survey and 200 teachers were selected for the telephone survey. Both designs were implemented using sampling without replacement.

Instrument response rates. Teacher response rates were calculated by dividing the number of teachers agreeing to take part in the study by the intended number of sampled teachers in that stratum. As shown in Table 5 the response rate for the telephone survey was considerably higher than the response rate for the written survey. This is a typical finding in nearly all survey research studies. With the exception of the large-district-low-scoring and the small-district-no-CSAP strata, the response rate for the written survey was at the high end of what can be expected for mailed surveys.

Table 5

Instrument response rates per stratum.

Stratum	Written Survey Teacher response rate per stratum	Telephone Survey Teacher response rate per stratum
1 (Large/High)	46.6%	72.7%
2 (Large/Low)	18.2%	90.5%
3 (Med/High)	49.7%	78.3%
4 (Med/Mod)	44.7%	83.3%
5 (Med/Low)	41.0%	94.4%
6 (Small/High)	44.6%	63.2%
7 (Small/Mod)	51.9%	85.0%
8 (Small/Low)	50.0%	81.3%
9 (Small/NA)	25.0%	100%

Individual stratum response rates are aggregated in Tables 6, 7, 8, and 9 to provide a quick overview of how response rates varied in relation to the two organizing variables of district size and CSAP achievement level. For example in Table 6, it is apparent that teacher response rate on the written survey was poorer in districts with low CSAP scores, 36% as compared to 48% in the other two strata. Similarly, in Table 8, the teacher response rate to the telephone survey was poorer, 64% in low scoring districts as compared to 84% and 93% in the other strata. When teacher response rate is examined by size of district, it is clear that large districts had poorer response rates on the written survey. However, size of district did not have an effect on response rates for the telephone survey, where teacher response rates were high for all three strata, 80%, 83%, and 78% respectively.

Table 6

Written survey response rates for district teachers with High, Moderate, and Low CSAP scores.

Type of district by CSAP score	Total number sampled (written)	Response rate (written)
High	380	47.6%
Moderate	195	47.7%
Low	226	35.8%

Table 7

Written survey response rates for Large, Medium, and Small district teachers.

Type of district by size	Total number sampled (written)	Response rate (written)
Large	208	36.1%
Medium	372	46.2%
Small	229	48.0%

Table 8

Telephone survey response rates for district teachers with High, Moderate, and Low CSAP scores.

Type of district by CSAP score	Total number sampled (telephone)	Response rate (telephone)
High	98	92.9%
Moderate	44	84.1%
Low	77	63.6%

Table 9

Telephone survey response rates for Large, Medium, and Small district teachers.

Type of district by size	Total number sampled (telephone)	Response rate (telephone)
Large	54	79.6%
Medium	88	83.0%
Small	58	77.6%

Standard errors and design effects. Due to the stratified cluster survey sampling design, standard error (SE) estimation could not be completed using a simple random sampling model. Statistical software packages such as SPSS would underestimate the SEs because they assume random sampling and would ignore the effect of clustering by district. Therefore, the "proc surveymeans" procedure through SAS was used to compute standard errors. SAS software utilizes the Taylor expansion method to accurately compute errors. This procedure estimates variance through the primary sampling units (districts) due to the effect of clustering.

Data analyses were carried out to obtain basic statistical measures such as means and standard deviations, frequencies, proportions, and standard errors. For these analyses, each participant response was valued as follows:

$$\overline{\hat{Y}} = \left(\sum_{h=1}^{H} \sum_{i=1}^{I} \sum_{j=1}^{J} w_{hij}\, y_{hij} \right) \Big/ w...$$

h= 1, 2, ..., H is the stratum number, with a total of H=9 strata
i= 1, 2, ..., n_h is the cluster number within stratum h, with a total of n_h clusters
j= 1,2, ... , m_{hi} is the unit number within cluster I of stratum h, with a total of m_{hi} units
y_{hij}= observed values for observation j in cluster I for stratum h
w_{hij}= sampling weight for observation j in cluster I for stratum h
$w...$= the sum of the weights over all sample observations

See Weighting section under Data Processing for further weighting details.

Field Operation Setup and Data Collection

Pilot testing. The questionnaire was developed and revised through an iterative process, based on piloting of questions with 15 randomly sampled teachers across Colorado in December 2000. Pilot questions were open-ended, and all pilot administrations took between 15 to 20 minutes via telephone. Originally, the questionnaire was designed to be conducted only as a telephone survey. However, it was determined that a more representative sample with a larger number of respondents could be obtained if a written survey was used in addition to the telephone survey. The final questionnaires (both written and telephone) were based in part on the pilot test results.

The written and telephone questionnaires were grounded in the described pilot testing. Surveys designed by Kennedy, Ball & McDiarmid (1993), Porter (1994) as well as work done by CRESST researcher Brian Stecher were used as models for the two questionnaires (Stecher, B. M., Barron, S.L., Chun, T. and Ross, K. 2000). Ravay Snow-Renner, a senior research associate at McREL, also assisted in the design of the instruments.

Written questionnaire. The surveys (see Appendix A) were sent in packets including the questionnaire, a stamped, self-addressed envelope, a stamped, self-addressed postcard used for tracking purposes, and a pen. The packets were mailed between March 16 and March 20, 2001, and all responses were received between March 20 and June 12, 2001.

Of the 809 sampled participants, there were a total of 357 respondents (44% response rate) for the written questionnaire. This response rate is considered average to high (response rates for mailed surveys are typically between 10% and 50%), perhaps due to the timely and politically relevant nature of the survey questions (Weisberg, Krosnick,

& Bowen, 1996). Three hundred sixty-four teachers returned questionnaires; however, seven surveys were discarded because the respondents did not include information needed to ascertain cluster data.

Study weights were applied to the available data in each stratum to preserve original population proportions regardless of the response rate per stratum. This procedure ensures that teachers from districts of each size and type are represented in proportion to their numbers in the total population. Note that this statistical weighting is not intended to correct for response biases caused by differences between teachers who chose to respond and those who did not. The much higher response rate of the telephone survey serves as a check on potential respondent bias in the mailed survey results.

Phone questionnaire. Person-to-person contact was made with all sampled telephone survey participants except for one individual. Teachers who agreed to take part in this survey spent 20 to 30 minutes responding to the interview protocol as found in Appendix B.

Of the 200 sampled teachers, there were a total of 161 respondents (80.5% response rate) who agreed to participate. This is a high response rate compared to a national telephone survey completion rate of approximately 60% (Weisberg, Krosnick, & Bowen, 1996). Again, study weights were applied to the available data in each stratum to preserve original population proportions regardless of the response rate per stratum.

Data Processing

Telephone open-ended response data were transcribed subsequent to each phone interview. Quantitative telephone data were entered into an SPSS spreadsheet by survey staff. Written questionnaire free-response data were transcribed by survey staff, and quantitative data were saved as an Excel spreadsheet and in ASCII format by Mail Graphics, Inc.

Qualitative coding. The same process was used for coding qualitative data from both the phone and written questionnaire. The coding process was completed in five stages:

Stage 1: 3-5 survey staff members independently looked for themes in all of the transcribed responses for a given question;

Stage 2: After independently determining themes within question responses, staff members convened for a coding meeting to determine common themes across coders;

Stage 3: As determined at the coding meeting, the most consistently identified themes became the codes for each question;

Stage 4: Two survey staff members independently coded question responses based on the aforementioned codes;

Stage 5: If the staff members were in agreement on question response codes, then the response was officially coded upon agreement. If staff members disagreed on a response code, the response was read by a third survey staff member, and a final decision was jointly agreed to by the readers.

Please see Appendix C for the qualitative data coding schemes.

Quantitative processing.

Instrument cleaning and quantitative coding. Once data entry was complete, we examined the data for quality by spot-checking with the original surveys. For teacher type, "Other," we recoded the teacher types as follows:

- Specialist: Title I, special education, reading specialist, resource room, and ESL teachers;
- Elective: Physical education, typing, drama, business, foreign language, shop, agriculture, home economics, debate, art, journalism, music, and technology teachers;
- Gifted and Talented: Both elementary and secondary gifted and talented teachers;
- Interdisciplinary teacher: Teaches two or more subjects for four or more grades;
- Non-classroom teacher: Counselors, speech therapists, occupational therapists, etc.

For "Gifted and Talented," "Interdisciplinary," and "Non-classroom teacher," the frequencies of responses were quite low. Therefore, these data were not usable for reporting cross-tabulated responses.

Other recoded variables were as follows:

- Because teachers were asked about grade level but not what type of school they worked in, school type was coded as:
  - Elementary: Teaches within grades K-5.
  - Middle School: Teaches within grades 6-8.
  - High School: Teaches within grades 9-12.
  - Mixed: Teaches across Elementary/Middle School/High School grade levels.
- "Valid" respondents were those whose questionnaires were returned with district identification data included.

Weighting. Because the participants were not chosen through simple random sampling, there was not equal chance of all teachers in Colorado being picked. Therefore, all respondents' answers were weighted based on the probability of each teacher being in this study.

$$z_{hij} = \frac{N_h}{N} \times \frac{N_{hi}}{n_h} \times \frac{1}{n_{hi}}$$

$$w_{hij} = (z_{hij}) \times (\text{Total teacher population})$$

z_{hij}= Probability of a given respondent (j) in a particular stratum (h) and cluster (i) being in this study
N= Total teacher population for Colorado
N_h= Total teacher population for stratum (h)
N_{hi}= Total teacher population for cluster (i) within stratum (h)
n_h= Sampled teacher population for stratum (h)
n_{hi}= Total respondents for cluster (i) within stratum (h)
w_{hij}= Respondent weight within total teacher population

Using these terms, each probability within the z_{hij} calculation was determined:

$\frac{N_h}{N}$ = p_h= Probability of respondent coming from stratum (h)

$\frac{N_{hi}}{n_h}$ =p_{hi}= Probability of respondent coming from cluster (i) within stratum (h)

$\frac{1}{n_{hi}}$ =p_{hij}= Probability of being a respondent from cluster (i) within stratum (h)

By multiplying the probabilities, we determined the joint probability (z_{hij}) of a given respondent (j) in a particular stratum (h) and cluster (i) being in this study.

Weights and all statistics were computed separately for telephone and written questionnaire data.

Findings

Standards-Based Reform

Given that content standards were established as a mechanism to improve schools, teachers were asked what they thought of standards. In open-ended questions on both the written and phone survey, teachers voiced generally positive feelings about standards. On the phone survey, when asked if they had made any changes in their classroom instruction because of the standards, 86.9% teachers replied "yes" and 13.1% said "no". The two most frequent changes teachers reported were *aligning* curriculum with standards and *adding* something to the curriculum because of standards. Quotations typifying these two themes are listed below in descending order of prominence (All open-ended question theme frequencies can be found in Appendix D).

• Teachers reported aligning curriculum/instruction/lessons with standards.

Once the state content standards were released, school districts went through the process of aligning their curriculum with the standards. To varying degrees, teachers also aligned their classroom instruction to the standards.

"I'm the chair of the science department. We sat down and went through the curriculum for the various science classes and we made sure the curriculum was aligned with the Colorado Science Standards. We made sure there was one-to-one correspondence between the two. We covered the materials that are required by Colorado science standards and then we build upon that, added to it. We've done that with all our classes. Individually, teachers may emphasize one area more than another, but relative emphasis is not specified in the Colorado standard. The basic material is covered."

"We have gone through and aligned all the standards for fifth and six grades so I know in fifth grade what I should be teaching. Within the last two years it constitutes a change I would say. We also have standard-based report cards at our school. I would say that since we have divided standards by grade levels, I am a lot more aware of what I am teaching that matches what the standards want. I try to go out of my way to touch all the standards during the day."

"Standards-based reform has forced me to not do lessons in isolation and to always try to cross integrate subject matter. For geography, it is history, civics, and economics and geography itself. So I'm always trying to take those standards from those four strands and weave them together into complex events, discussions, field work and surveying... So since they are spelled out specifically, the only way to combine them is into complex learning environments."

•Teachers reported adding content to their curriculum as a result of standards-based reform.

Teachers reported adding a variety of content topics to their curriculum. For example, in math, teachers noted that they had added probability, statistics, and more

geometry to their curricula because of standards and were using TI 83 graphing calculators with greater frequency. Teachers in other content areas echoed math teachers' reflections.

"Number one, in math, I am doing more concept-oriented problems rather than calculation. So I'm doing more things with graphs and more things with probability, so that is pretty big in math. In reading more extended response answers to things that they have read and I'm doing that same thing, extended response written answers in math also. Those are the two big things."

"I added a nine week unit on the Middle East based on the standards. My lessons are geared more to the standards. I now include free-hand map making that is part of the standards. I did not do that before."

'I'm teaching things I didn't really focus on quite as much before like gathering information from graphs, charts and tables. We're doing a lot more in geometry and calculators, story problems and with sheer number sense as well as computation skills."

As is illustrated by the foregoing responses from the telephone survey, the majority of teachers reported shifting their curricula to meet state content standards. In the next section, survey results comparing standards based reform and CSAP testing are discussed.

Standards-Based Reform and CSAP

"You know, I looked at what the standards were and looked at what I was doing, how it fit. So, I modified it some, but I don't think the real drastic changes came until the release of the CSAP test."

CSAP was developed in the context of standards-based reform. Both attention to CSAP and implementation of standards would be expected to affect changes in curriculum and instruction. Therefore, parallel sets of questions were asked about the impetus for various changes.

On the written survey, teachers were asked to rate the extent to which standards-based reforms and CSAP testing have led to changes in a variety of factors that affect the quality of schooling: professional development, quality of classroom instruction, teaching higher-order thinking skills, faculty morale, attention to students at different levels of performance, student access to elective classes, amount of emphasis placed on writing and use of multidisciplinary approaches to subject matter. Teachers were first asked to rate whether each factor had increased, stayed the same, or decreased due to standards-based reform (SBR). Then teachers were asked to rate each feature in terms of the impact of CSAP testing. Data in Table 10 reflect the extent to which teachers attributed changes in schools to both standards and CSAP testing. Percentages, means, standard deviations and standard errors are shown. Responses were coded as follows: "Increased" as 3, "No change" as 2, and "Decreased" as 1.

BEST COPY AVAILABLE

Table 10

Written survey question: ***To what extent have standards-based reform and CSAP testing led to changes in the following areas?***

	Standards-Based Reform (SBR)			CSAP		
	Decreased	No Change	Increased	Decreased	No Change	Increased
Valuable professional development opportunities SBR X= 2.49 (s.d. .031) CSAP X= 2.18 (s.d. .032)	4% Freq.=13 SE=.009	43% Freq.=145 SE=.032	53% Freq.=182 SE=.030	14% Freq.=48 SE=.018	55% Freq.=179 SE=.023	32% Freq.= 106 SE=.021
Quality of classroom instruction SBR X= 2.52 (.023) CSAP X= 1.99(.038)	4% Freq.=16 SE=.006	40% Freq.=129 SE=.022	56% Freq.=196 SE=.022	29% Freq.=101 SE=.030	43% Freq.=142 SE=.039	28% Freq.=91 SE=.024
Teaching higher-order thinking skills SBR X= 2.40(.036) CSAP X= 2.07 (.047)	8% Freq.=28 SE=.014	44% Freq.=146 SE=.029	48% Freq.=164 SE=.030	24% Freq.=82 SE=.027	44% Freq.=143 SE=.028	32% Freq.=103 SE=.028
Faculty morale SBR X= 1.76 (.043) CSAP X= 1.21 (.029)	35% Freq.=122 SE=.030	53% Freq.=181 SE=.028	12% Freq.=40 SE=.020	81% Freq.=273 SE=.029	18% Freq.=59 SE=.030	1% Freq.=4 SE=.005
Attention to lowest performing students SBR X= 2.42 (.029) CSAP X= 2.34 (.041)	5% Freq.=17 SE=.009	48% Freq.=160 SE=.036	47% Freq.=166 SE=.031	12% Freq.=43 SE=.024	42% Freq.=140 SE=.035	46% Freq.=155 SE=.029
Attention to students in the middle range SBR X = 2.27 (.028) CSAP X=2.29 (.031)	5% Freq.=17 SE=.011	62% Freq.=212 SE=.024	33% Freq.=113 SE=.024	8% Freq.=28 SE=.011	55% Freq.=184 SE=.030	37% Freq.=126 SE=.028
Attention to highest performing students SBR X= 2.11 (.033) CSAP X= 1.97 (.026)	14% Freq.=45 SE=.015	62% Freq.=216 SE=.031	25% Freq.=81 SE=.028	23% Freq.=79 SE=.023	57% Freq.=195 SE=.037	20% Freq.=64 SE=.022

Student access to elective classes such as art, music and physical education SBR X = 1.84 (.029) CSAP X = 1.74 (.033)	21% Freq.=71 SE=.028	75% Freq.=255 SE=.030	4% Freq.=13 SE=.010	27% Freq.=97 SE=.033	71% Freq.=232 SE=.034	2% Freq.=5 SE=.006
Amount of emphasis placed on writing SBR X= 2.75 (.031) CSAP X= 2.75 (.042)	0% Freq.=1 SE=.003	24% Freq.=86 SE=.030	75% Freq.=256 SE=.030	4% Freq.=12 SE=.014	17% Freq.=58 SE=.024	79% Freq.=266 SE=.032
Use of multidisciplinary approaches to subject matter SBR X =2.26 (.033) CSAP X = 1.95 (.056)	14% Freq.=49 SE=.020	47% Freq.=157 SE=.026	39% Freq.=138 SE=.022	29% Freq.=97 SE=.037	48% Freq.=156 SE=.030	24% Freq.=84 SE=.025

Across many of the features, teachers reported more positive changes due to standards than to CSAP testing. This pattern can be seen by comparing means or the percentages reporting positive increases. The pattern of standards leading to more positive changes was consistent across school levels (elementary, middle school, and high school), subject areas, and overall self-reported levels of school performance (see Appendix E).

The difference between standards-based reform and CSAP testing was most pronounced for faculty morale, which showed a decrease due to standards but a more substantial decrease due to CSAP testing. 81% of teachers reported a decrease in faculty morale because of CSAP.

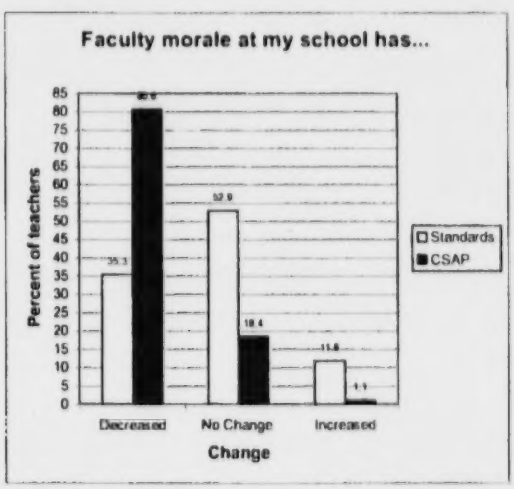


Figure 1. Change in faculty morale due to SBR and CSAP testing.

Additionally, 56% of teachers reported that the quality of classroom instruction had improved with standards-based reform while only 28% said that it had improved with CSAP testing.

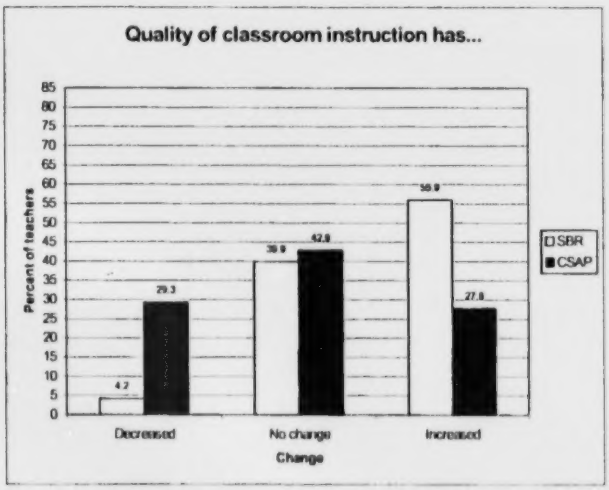


Figure 2. Change in quality of classroom instruction due to SBR and CSAP testing.

Teachers' ratings of both standards-based reforms and CSAP testing were more similar for the amount of emphasis placed on writing, with teachers reporting substantial increases in the amount of writing for both standards-based reforms and CSAP testing.

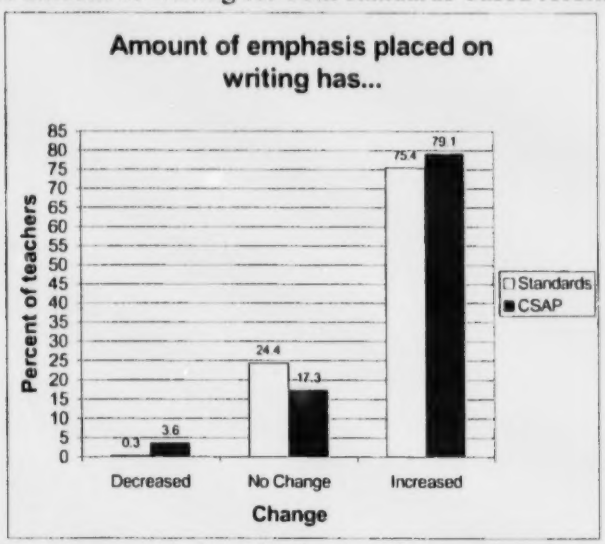


Figure 3. Change in amount of emphasis placed on writing due to SBR and CSAP testing.

Teachers' responses on the phone survey followed the same overall pattern as the written survey data (see Appendix F). Specifically, the phone survey responses reflected a tendency for teachers to rate standards-based reform as having "increased" reform-

based instructional practices and CSAP as having made "no change." The only difference was that on the written survey, teachers reported more opportunities for professional development because of standards than because of CSAP. However, on the phone survey teachers indicated that professional development opportunities had increased under both standards-based reforms and CSAP testing.

In pilot interviews conducted with fifteen teachers before we designed the surveys, teachers had reported decreased student access to elective classes such as art, music and physical education due CSAP testing. However, only 27% of the teachers in the more representative written survey reported decreased student access to electives because of CSAP. In the phone survey only 22% reported a decrease in student access to electives.

Teachers' perceptions of professional development opportunities available with standards-based reform vs. CSAP varied by level of school performance. While overall, 53% of teachers in the written survey reported that valuable professional development opportunities increased under standards-based reform in contrast to only 32% of the teachers reported an increase with CSAP, the percentages looked very different when teachers were divided into groups by level of school performance.

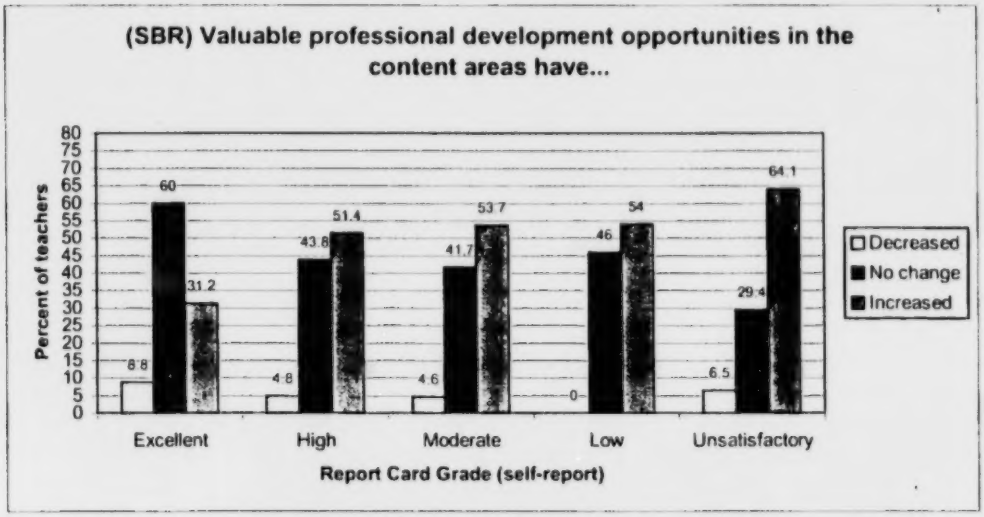


Figure 4. Change in the amount of valuable professional development opportunities due to standards-based reforms.

As illustrated in Figure 4, 64% of teachers who reported that they worked in schools that would be rated "unsatisfactory" on the state school report card believed that opportunities for valuable professional development had increased due to standards-based reforms. In contrast, only 31% of teachers at "excellent" schools reported an increase in valuable professional development opportunities. Simultaneously, 24% of teachers at "unsatisfactory" schools reported a decrease in valuable professional development opportunities under CSAP as illustrated in Figure 5.

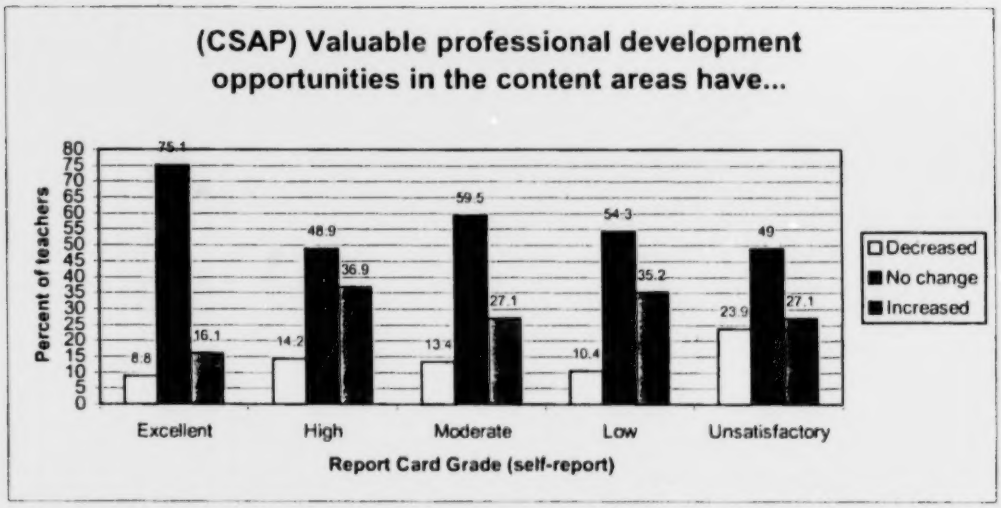


Figure 5. Change in the amount of valuable professional development opportunities due to CSAP testing.

The three most salient themes to emerge from the data comparing standards-based reform and CSAP testing are listed here in descending order of prominence:

• **Teachers reported a decrease in morale due to CSAP testing.**

In both open-ended and closed-ended questions on the written and phone survey, teachers consistently reported not only that their morale had decreased because of CSAP testing but also that school morale overall had declined.

"We are under so much pressure to get good results there is little time for anything else. CSAP definitely decreases morale."

"The CSAP did not cause me to improve my classroom. It caused me to leave the classroom. My students always did well on the ITBS tests, but when there was so much pressure and more stress on how we would do as a school, I decided to teach art and leave the classroom. It went against my grain and my heart to teach to a test. I believed in familiarizing my students with the format and having them work to my high standards and to making them accountable, but I could not compromise my beliefs and teach to the test. Teaching is tough enough when you work in a transient community in rural Colorado without someone from the government threatening to change you over to a charter school. The test itself is not a bad test. The way the results are used is absolutely horrible! Issuing school report cards is very degrading to those in the education field. Some of the best teachers have some of the lowest kids."

"I had to cease a lot of projects and other activities & programs in my room to make time for the teaching of test-taking strategies. I felt demoralized and unappreciated by all of the negative press in the newspapers and have doubted myself as an educator for the first time. I'm not sure I would go into this profession if I had to choose all over again. I feel pulled from many directions - to make education more personal and then,

from the CSAP - to standardize their learning - forcing them into a box whether they are ready developmentally or not."

• Teachers believe that standards- not CSAP testing- have improved the quality of instruction in their classrooms.

Open ended questions on the written survey provided greater opportunities to learn teachers' rationales for rating the influence of CSAP testing differently than standards. When teachers were asked to describe one or two examples of how CSAP had helped improve the quality of instruction in their classroom, teachers on the phone survey asserted that standards, not CSAP tests, had helped them improve the quality of instruction in their classroom.

"I believe that state standards, not CSAP, have improved the quality of instruction. My teaching is more focused on district benchmarks."

"My quality of instruction has improved due to standards-based criteria and workable rubrics and assessments not because of CSAP."

"It [CSAP] hasn't improved instruction, just changed it. We now teach to CSAP."

"If anything, CSAP takes time away from critical thinking, research, and discussion because it requires some format preparation and then administration which takes the better part of two weeks away from instruction."

• Teachers reported an increased school-wide emphasis on writing due to both CSAP testing and standards-based reforms.

Teachers noted a greater school-wide emphasis on writing. Math, Science, Art, and Physical Education teachers reported that they were emphasizing writing more in their classroom, and many had been trained in commercial programs such as Six-Trait Writing or Step up to Writing.

Math teacher: *"We are writing more school-wide using Six-Trait Writing. I have students write out explanations for problems. Every homework assignment, every test has several examples on it. The kids really got good at doing it the way I think the test-makers wanted them to write out the response."*

Science teacher: *"Rather than focusing on content in science, we are focusing on literacy and more reading and writing. We have been doing things like free-response. We have been doing two column notes a lot."*

Art teacher: *"The whole school has been trained in Step up to Writing. In art, I have the students read articles from Scholastic Magazine on artists. I have them do more writing. They have to describe the process involved in creating a ceramic pot."*

Physical Education teacher: *"We did Six-Trait Writing. My class has to do three or four writing assignments per quarter to demonstrate writing skills . . . I asked them to*

write about teamwork before. If I have a class that shows good sportsmanship, I have them write about sportsmanship and things like that."

As is illustrated by teachers' responses on both the written and phone surveys, standards-based reform and CSAP testing have affected Colorado's teachers. Standards-based reform was viewed more positively than CSAP testing; teachers believed that standards have helped them improve the quality of classroom instruction. Teachers credit both standards and CSAP testing with pushing them to emphasize writing more in their classroom.

In the next section, survey results concerning the influence of CSAP on instructional practices will be discussed.

The Influence of CSAP Testing on Instruction

There is a substantial body of research suggesting that high-stakes assessments influence instructional practices (Popham, 1987; Darling-Hammond and Wise, 1985; Rottenberg and Smith, 1990). For example, teachers may add or eliminate curricula to align instruction more closely with the assessment. Additionally, test preparation activities can significantly change the instruction children receive in a school year to the extent that test preparation activities supplants normal instruction (Shepard and Dougherty, 1991). Finally, teachers may target specific groups of students for additional help to improve overall test scores in a class or school. In addition to a "global" set of questions comparing the effect of standards versus CSAP testing, the survey instruments also included more detailed questions about both positively and negatively perceived effects of testing on instruction.

Aligning instruction with CSAP. In both the phone and written surveys, teachers reported making changes to align instruction with CSAP. In some cases this meant adding content, such as more reading and writing instruction. In other cases, it meant eliminating subject matter such as science and social studies. While teachers reported emphasizing problem solving more and asking students to explain their reasoning, teachers also reported reducing the number of labs, field trips, and extended projects.

In the phone interview when teachers were asked if CSAP testing had changed the content of instruction in their classroom, 62.7% said "yes" and 37.3% said "no." Out of the 100 teachers who responded affirmatively, 86% reported that they had added something to their curriculum to align it with CSAP. 82% of the teachers reported eliminating something from their curriculum to align it with CSAP.

Similarly, in the written survey, the most common response from teachers regarding CSAP was that they had either added or eliminated curricula to make their instruction more aligned with CSAP. The most salient themes to emerge from the written and phone survey related to specific instructional practices are listed below in descending order of prominence:

• **Teachers reported adding something to their curriculum to align it more closely with CSAP.**

Just as school districts and teachers went through a process of aligning curriculum to standards, some are now gradually going through a process of aligning their curriculum to CSAP.

"Now instead of just teaching earth and life science to the grade level, I have to pick up the physical sciences. We are hitting more strongly on the elements and compounds."

"I included this year a geometry exam and asked students to reflect on the procedures they followed and explain them because that is something required by the test. That is something important that I need to include more in my instruction. I am also going to include the theme of probability because that is something included in the test too. I am also going to include tessellation."

"I think it gives me justification for doing more writing in the math classroom, the whole push for being able to explain their reasoning, I think this wasn't happening across disciplines through the classes before CSAP came along. I probably do it more because I feel more secure and have more support for doing it now. I think it gives me more support for doing problem solving in the classroom, instead of just routine work."

• **Teachers reported emphasizing reading and writing more in their classrooms.**

In an effort to align their instruction with CSAP, teachers reported emphasizing writing and reading more with their students.

"Yes, for example, with reading curriculum that I teach it has become much more focused and I think that is one of the good things about CSAP is that I know or we know what the right requirements are now on CSAP. I feel like my writing instruction has greatly improved over the years as it should but I also think that CSAP has played a role in that because it has given me a much better focus on writing and what needs to be taught. So I think that is a good thing. And it is the same with reading as well. You get into the dilemma of teaching to the test. I feel that the things that I teach in reading and writing are good things to know and now are they teaching to the test? Maybe you could say it is but is it still valid, good information that kids need to know? Yes. Is it curriculum based? Yes. So, it has just made me much more focused and I think it has made me a better teacher."

"I feel I've really improved in my instruction of writing. The majority of my students feel confident in writing an accordion paragraph. They also do a nice job with the various text structures—compare/contrast; description; problem/solution, etc."

"Writing. We write a lot more. A lot more personal writing, expository. Paragraphs. Before we did not write as much . We do 6-Trait a lot more. Reading

hasn't really changed except that we practice to the test a lot more. But writing has changed a lot. Reading, yes I read a greater variety of genres. We read a lot more nonfiction."

•Teachers reported either eliminating or cutting back on the amount of science and social studies they teach.

Across grade levels, teachers reported reducing the amount of time they spent teaching science and social studies. In some cases, this was because school and district officials had advised them to do so. In other cases it was a choice made by individual teachers.

"I eliminated a lot of my social studies and science. I eliminated Colorado History. What else? Electricity. Most of that because it's more stressed that the kids know the reading and the math, so, it was pretty much said, you know, do what you gotta do."

"Our district has told us to focus on reading, writing, and mathematics. Therefore, science and social studies, unless I can teach them in a reading, writing, or mathematics format then they don't get taught. I don't teach science and social studies nearly as often and not purely as science or social studies. In the past I had hatched out baby chicks in the classroom as part of the science unit. I don't have time to do that. I have dissected body parts and I don't have time to do that. I have waited until now to start my economics unit for mini-society because it takes so much time. I can do it now because CSAP is over. We don't take as many field trips. We don't do community outreach like we used to like visiting the nursing home or cleaning up the park because we had adopted a park and that was our job was to keep it clean. Well, we don't have time for that any more."

"Those things (science and social studies) just fall to the back burner and quite frankly, I just marked report cards for the 3rd grading period and I didn't do science at all for the third grading period. Same for the social studies."

"I know part of the science and social studies has been cut back to give more time to those other areas. The units themselves were cut and we haven't been able to meet all the standards. We couldn't do weather and climate because of lack of time."

•Teachers reported emphasizing problem-solving more in their instruction.

"I focus on students' ability to explain their problem solving more than just finding the right answer."

"I have included more reading of articles including scientific articles in my class as well as story type/long problems for problem solving. It reminds me that it can not all be fact based."

"Everything they do has to be thoroughly explained. For example, on math, they have to explain their reasoning behind it and give their words. The program we have

requires that... they have [to] give examples depending on how the CSAP is set up. And we do number talks. Anything from example problems from old CSAPs to just talking about things. Going through problems as a group. Going through thinking skills strategies... I use some of the Marcy Cook things, some of the stuff she has come out with... just things like that."

• Teachers reported eliminating instructional formats such as lab work, projects, and field trips.

Because of increased emphasis on reading and writing as well as test preparation activities, many teachers reported less time for activities such as lab work, research projects and field trips.

Middle School Science Teacher - *"I had to cut out some things in order to do the CSAP stuff. It's not the number of days. I think it would be more accurate to say the number of labs. I think what is more significant is that I have had to cut the number of hands-on investigations. I would say I have had to cut one quarter of the labs."*

"We only teach to the test even at 2nd grade, and have stopped teaching science and social studies. We don't have assemblies, take few fieldtrips, or have musical productions at grade levels. We even hesitate to ever show a video. Our 2nd graders have no recess except for 20 minutes at lunch."

"Anthology- had to stop working on it. We also eliminated the long term month to month projects that involved writing as well."

"As a social studies teacher, I had to give up many content days that might have been spent on projects or discussion activities that were spent reinforcing what they were learning in English. A great deal of time was wasted on test-performance skills."

"Projects, eliminated curriculum such as novels I would teach, we didn't have time to go to the library, we didn't have time to use the computer labs, because they had to cut something. Cut things I thought we could live with out. Cut presentations, anything that takes very much time, I cut film. We have been cutting like crazy."

In order to align their curricula more closely with CSAP, teachers made difficult choices as to what to cover and what to leave out. In many cases this meant emphasizing writing and reading at the expense of science and social studies. Long term projects and presentations had to be given up as well. However, as we discuss next, teachers' decisions about what to teach and what to leave out varied considerably according to their school's performance level. Teachers who self reported that they worked at schools that would be classified as "excellent" by a school report card made very different decisions than teachers who worked at schools considered to be "unsatisfactory."

Differential impact by level of school performance. In the written survey, we asked teachers to rate how often they engaged in a variety of instructional practices (see

BEST COPY AVAILABLE

Appendix G). The questionnaire items were designed to include both traditional practices focused more on basic knowledge as well as reform-based practices aimed at more challenging reasoning, problem-solving, and applications activities. As shown in Appendix G, the most frequent instructional practices across the state are demonstration of knowledge of basic skills and vocabulary, reported to be used frequently by 91% of teachers, as wells as application of concepts and principles, used frequently by 84% of teachers. In general, more traditional instructional exercises, such as practicing facts or procedures, are used frequently by a greater percentage of teachers (69%) than more reform-oriented exercises, such as writing about how to solve a problem (48%). But, in fact many teachers use a combination of these strategies.

Secondly, we asked teachers if the frequency of these practices changed in preparation for CSAP testing (see Appendix H). As shown in Appendix H, most teachers did not report changing the listed instructional practices in preparation for CSAP testing. By setting up a contrast in the written questionnaire between instructional practices used throughout the school year and those used only in preparation for CSAP testing, we found fewer changes attributed to CSAP than had been reported in telephone interviews. After having made more pervasive changes in their curriculum and instruction in response to standards and CSAP, the majority of teachers did not report making any (short term) changes "in preparation for CSAP testing." Consistent with other data showing an increased emphasis on writing due to CSAP, 36.5% of teachers reported an increase in the use of class time to discuss examples of good writing. Additionally, approximately one-third of teachers reported spending more time practicing facts and procedures such as grammatical rules and having students review basic skills or vocabulary in preparation for CSAP testing. Note that these shifts are increases in activities that already had high frequencies.

The largest decrease reported by teachers associated with CSAP testing was that twenty-two percent of the teachers reported decreases in conducting "projects that extend over several days" and "research in the library or on the internet."

When change data were broken down by anticipated school report card grade, some interesting patterns emerged. While 11.8% of teachers from "excellent" schools reported no changes in conducting projects that extend over several days, 42.3% of teachers from "unsatisfactory" schools reported doing them less often (see Figure 6). Thus, CSAP had more of an effect on instructional practices in low-performing schools, shifting attention more toward drill and practice on basic skills. Similarly, while 17.6% of teacher from "excellent" schools reported doing less research in the library or on the internet, 43% of teachers at "unsatisfactory" schools reported eliminating such student research projects (see Figure 7).

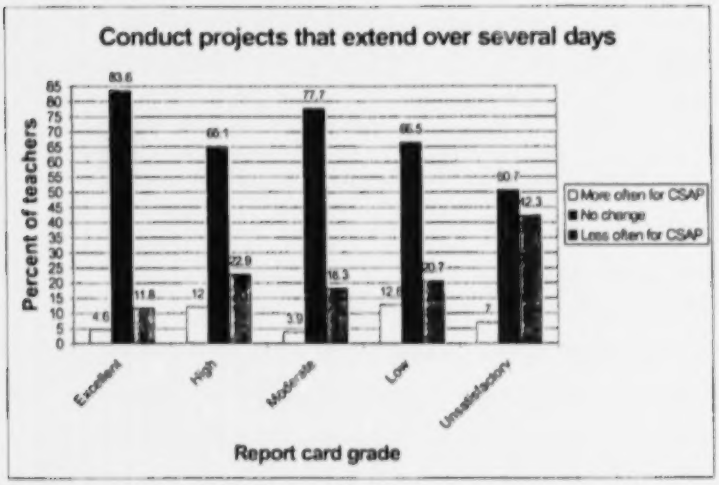


Figure 6. Frequency of conducting projects

And again, 17.6% of teachers at "excellent" schools reported a decline in using computers as part of regular instruction while 37.6% of teachers at "unsatisfactory" schools decreased the use of computers (Figure 8).

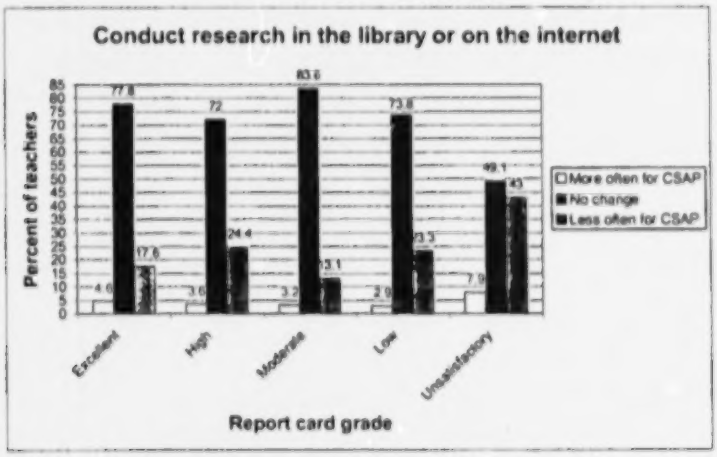


Figure 7. Frequency of conducting research.

BEST COPY AVAILABLE

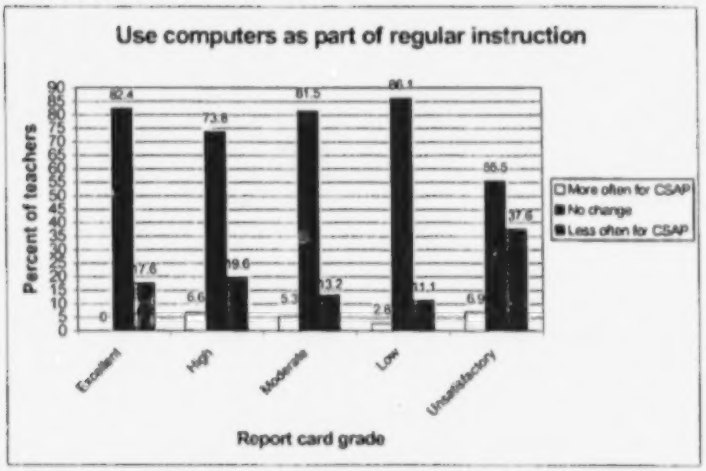


Figure 8. Frequency of using computers.

In writing, unsatisfactory schools were split (see Figure 9). 48.4% schools of the lowest-performing schools reported that they read and discussed examples of good writing more often while 16.3% reported doing it less often. In contrast, a majority of the schools at the excellent, high, moderate and low levels reported no change in the amount of writing. The change data should be interpreted in light of the general trend to increase focus on writing throughout the year in response to both standards and CSAP testing. Thus, for most schools there was not a change in attention to writing in preparation for CSAP testing perhaps because of a more pervasive focus on writing.

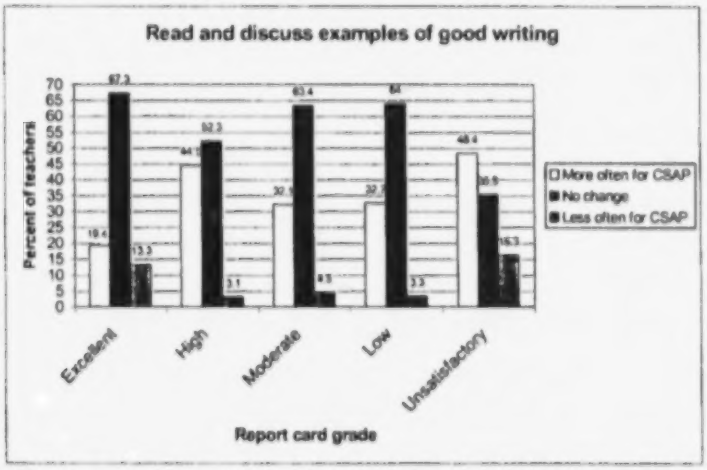


Figure 9. Frequency of reading and discussing examples of good writing.

Test preparation activities. Both the types of test preparation practices as well as the amount of time teachers spent on test preparation activities varied greatly from teacher to teacher. Overall, respondents frequently reported changing the format of instruction to make it similar to CSAP to help students feel more comfortable with the test. Additionally, elementary teachers reported using commercially produced test

preparation material as well as released items from the State to prepare students to take the test. A majority of teachers across all grade levels noted that they spent time on test-taking skills with their students.

In both the phone and written surveys, teachers were asked to report how much time they spent with their students doing test preparation activities. The amount of time spent on test preparation activities varied considerably by level of instruction (elementary, middle, or high school), and level of performance of the school (excellent, high, moderate, low, or unsatisfactory).

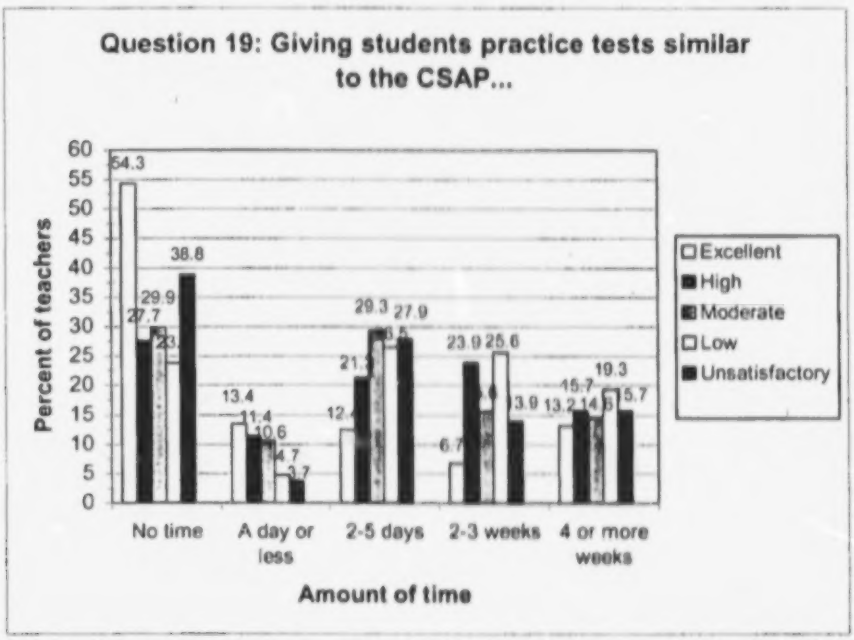


Figure 10. Amount of time spent giving students practice tests similar to the CSAP.

Teachers at different school performance levels spent disparate amounts of time on practice tests and test-taking strategies (see Figures 10 and 11). For example, teachers in "Excellent" schools reported spending less time giving their students CSAP practice tests than teachers in "High," "Moderate," "Low," or "Unsatisfactory" schools (see Figure 10). Although the amount of time spent taking practice tests varies consistently across all types of schools, ranging from "no time" to "4 or more weeks," spending two or more weeks is reported by 30-45% of teachers in the other four categories of schools but only by 20% of teachers in excellent schools.

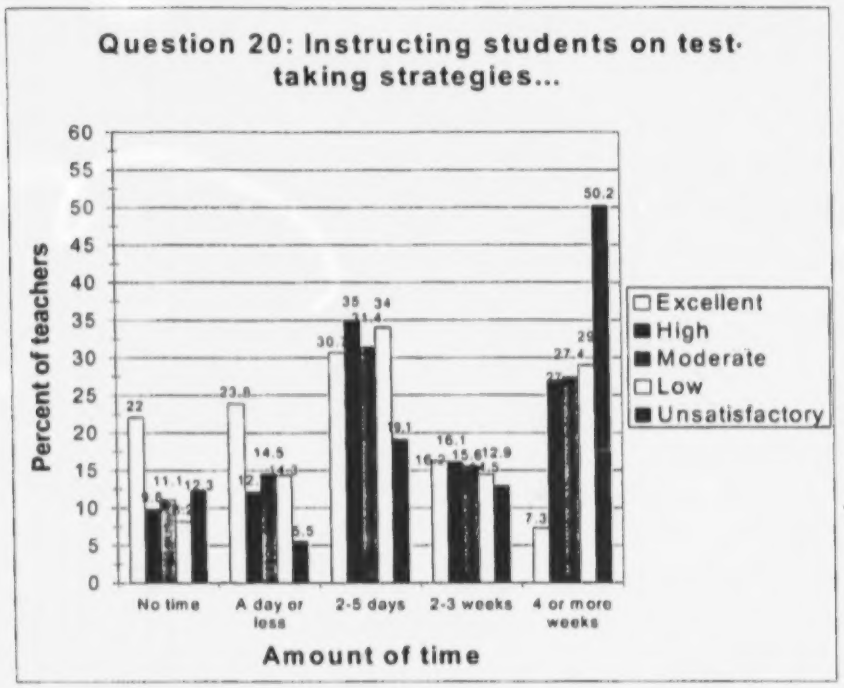


Figure 11. Amount of time spent on instructing students on test-taking strategies.

A similar pattern can be seen regarding time spent on test-taking strategies. 76.5% of teachers in "Excellent" schools spent 5 days or less instructing students in test-taking strategies. On the other side of the spectrum, 63.1% of teachers in "Unsatisfactory" schools spent 2 or more weeks on test-taking strategies. Approximately 57% of teachers from High, Moderate, and Low schools indicated that they spent 5 days or less on test-taking strategies (see Figure 11).

Differentiated time spent on CSAP preparation can also be seen in Figure 12. Note the disparity between the time that teachers at "Excellent" schools and "High," "Moderate," "Low," and "Unsatisfactory" schools prepare for CSAP testing. 43.8% of surveyed teachers at "Excellent" schools did not use CSAP prep activities or used such activities for only a few days before CSAP testing. 55% or more of teachers in the other four school categories reported that they spent time preparing for CSAP "regularly throughout the school year." Another 20-30% of teachers in these schools spent "a few weeks" preparing for the CSAP tests.

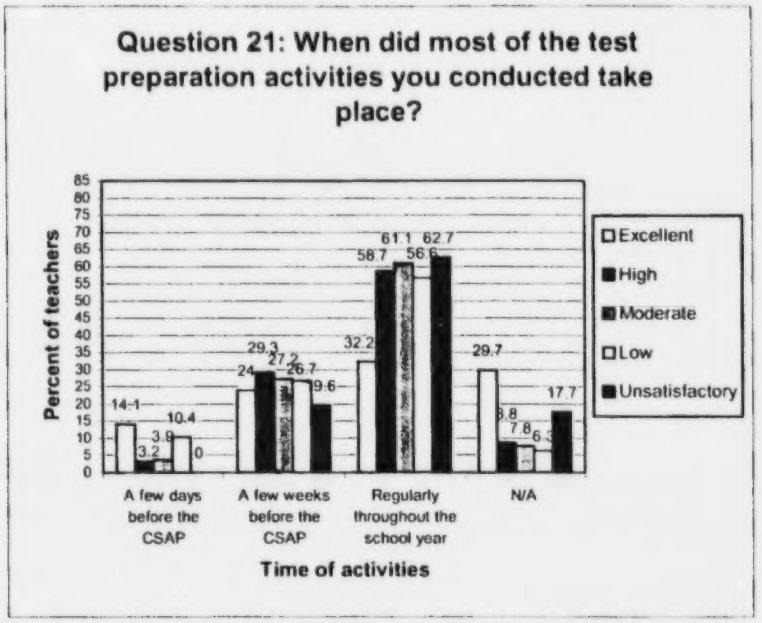


Figure 12. Test preparation activities.

In telephone interviews, teachers reported a range of test preparation activities. The most typical statements regarding test preparation are listed in descending order of prominence.

• Teachers included formats similar to the CSAP in their regular instruction to help their students be more comfortable with the CSAP.

"When I test children on the things that we have been learning, tests are given now with the little bubbles so that it's like a CSAP test, so every week we have a selection test for reading and it's given in CSAP format. I teach a lot more about inferences, we do a lot more summary, and my writing format has been much more extensive."

"We have changed the format of our daily oral language and we have retyped it ourselves because we couldn't find anything out in the market that would help us. So we have retyped it so that the format is similar to what the kids see on the CSAP. So that doesn't throw them so when they see that format and how to do something, they know how to go about it because it is not new and it is not strange."

"Well we're more aware of what the format is and the writing components and getting students to focus before they write, so we have the released items and we're aware of how they are organized and we're teaching that. It is not just vocabulary, its format, and of course we're hoping that they think though a problem and they don't just take the first answer that comes to mind, that they really give some thought to what the question is asking. Are they addressing the question? What we have found is that a lot of them just start writing and we are constantly saying, 'What was the question and what are you

supposed to be writing about?' You know, try to get that focus back. And then there were some checklists for kids, kind of like a rubric, a checklist for their writing, when they get it back, like, have they done this, have they put in the capitalization in when self-editing. We are much more in-tuned about that."

"There are in-services that go on and tell us what areas of reading and writing that the CSAP mostly tests on, key concepts, to make sure you cover, concepts to cover, just basically test-taking strategies, is what the CSAP is looking for, so that's what we teach to."

"In reading preparation for CSAP, throughout the year we read passages that we feel are very similar to a CSAP format, we definitely use the CSAP format in reading instruction throughout the year and we try to have the children focus on how to take a test more than on strictly reading for information and comprehension in it. I don't mean to sound like we are teaching totally to the test, but we want them to be very comfortable in that format. And so what we are doing is, we're having the children learn to read questions first and then having them read passages so that they are in tune for what we are looking for. . . .They want you to be able to compare genres in literacy, and be able to infer a lot about different non-fiction and fiction pieces. They want you to be able to predict certain ways, which a lot of it is good, because those are some really good strategies for reading.."

• Teachers used commercially produced test preparation materials similar to the CSAP and test items released by the State to provide their students with direct practice in preparation for CSAP.

"There is such a focus on preparing for the test. My district had copied several packets that I was obligated to go through and starting in January when we came back from Christmas, that was pretty much all I had to do was prepare for the CSAP.

There were reading packets that were the released sample items from the CSAP and there were math items that were thought to be similar to what would be on there. And we did those each day. That is how we would start our day. I have increased my math teaching time from 45 minutes a day to 90 minutes or more."

"With our Houghton-Mifflin series we have a set of 12 tests that they take so we get them used to taking a test and being able to show their work and explain good answers and things like that. So our reading series has geared them to taking tests. We have CSAP released items that practice with them to prepare them for the test. We do it pretty much all year."

"There is a whole bevy of commercial tests that you can procure. There is the old Iowa test and stuff like that. I bought it out of department money. They are produced by Exemplars and the other is called Test Ready. Then I have old copies of Iowa."

"In particular we spent a lot of time preparing-- meaning we bought the guides-- you know discarded things of previous years- and you know those things are very helpful

because children get accustomed to the format . . . All the way from late November to when we took the exam in February was pretty much . . . not the whole class time . . . but every part of each day was dedicated to doing practice passages and questions out of the CSAP materials. In preparation for CSAP, we pretty much just read the passage and you go back and you look at the questions.

I would try to do just one passage a day with some questions. Some time the passages were a couple of pages so we would continue it the next day. We have a 90 minute block. At least 30 to 40 minutes. So that is why everyone is thrilled now that no matter what our score is that at least it is over and we can do our thing. And it is not that we are deviating so much from standards or things that are important but at least now thematic units can take a part."

"We were instructed to ignore science and social studies in order to do test preparation CSAP booklets. We did this for roughly 2 to 3 months. In addition, the school taught a specific isolated skill each week."

• Teachers spent time instructing their students on test-taking strategies to prepare them for the CSAP.

Teachers reported devoting time to help their students with answer development: writing in complete sentences, organizing responses, answering all parts of the question, and writing on the lines.

"We do a lot of instruction with test-taking techniques, and give instruction on short written answers; the technique of short written answers, teaching them the techniques of short answers, of multiple choice, of long answers starting with a topic sentence, then add details and a conclusion, so that they can understand what is being asked of them in the question. In other words, we will look at questions about a certain excerpt we have just read, and analyze what the child is being asked to understand in the excerpt."

"I'm teaching more test-taking skills and how to use your time wisely. Also, what to look for in a piece of literature and how to underline important details. There is a lot more time spent on teaching those kind of skills. . . Read questions, restate the question in your answer, how to write so the person grading the test can read it, etc."

"Restating question in answer, going to text for answer; one reading questions; underlining; key words/concepts/found answers; using text as spellchecker; predicting answers based on titles, subtitles, etc. staying on the lines; proof reading work."

As demonstrated by both the written and telephone survey results, CSAP testing preparation activities influenced teachers' classroom instruction. The nature of one such instructional shift, "targeting" specific students, is discussed in the next section.

Targeting

BEST COPY AVAILABLE

In pilot interviews and discussions with school personnel, we had heard that to increase their test scores, some schools and districts were targeting students at the partially proficient level for extra help. As one district professional development employee explained, the district targeted "bubble kids." These were students who were on the cusp between partially proficient and proficient. To investigate this claim, we asked teachers in our survey first if they targeted any specific instructional efforts to help specific groups of students with the content knowledge measured by the CSAP? 68.6% teachers said "yes" and 31.4% said "no." For the teachers who had replied affirmatively, we asked them to describe their instructional efforts by proficiency level. Table 11 summarizes the results of the written survey for teachers who described differentiated targeting by levels of student proficiency.

Table 11

Targeting of Instructional Efforts Focused on Students at Specific Proficiency Levels (written survey, open-response)

Themes	Advanced	Proficient	Partially Proficient	Unsatisfactory
Instructional efforts focused on writing content	9	9	14	13
Instructional efforts focused on reading content	12	13	26	21
Instructional efforts focused on math content	4	2	2	2
Instructional efforts focused on higher level thinking content	24	7	2	0
Instructional efforts focused on basic skills and remediation content	1	0	4	4
Extra help (one-on-one, small groups, etc)	13	6	33	41
Total number of responses for each proficiency level	37	25	54	52

Although not as pronounced as in the pilot study, the written survey confirmed a trend for some teachers to give different amounts of attention to different groups. More attention was focused on partially proficient and unsatisfactory students.

When students were targeted by teachers, the most common intervention was additional assistance in reading and writing. Listed below are typical examples of what teachers said about the targeting of students.

•Teachers reported targeting students at the partially proficient level to improve overall CSAP scores.

5th grade Phone survey teacher from a rural school district

"We really target the kids in the middle because they can actually make a difference and change test scores. The high scoring kids are already high and you can't make as much of a difference as easily with the lowest performing kids."

3rd grade Phone survey teacher from a suburban school district

"All throughout the year we have targeted students who are on the cusp. We have developed ILP's for the students in order to help them bring up their reading and language abilities. We want to bring them up to where they should be and that indirectly affects the CSAP. It is not directed just at the CSAP, but at the whole child, so we have done a lot of things to bring them up to grade level. "

3rd grade phone survey teacher from an urban school district about partially proficient students

"Very strongly. Those are the kids that we pre-assess, probably on a bi-weekly, if not weekly basis. And again . . . the district has hired people to help with those specific kids. And we worked extremely closely together to help those kids starting in August. The person was brought on board in August. So, all I can say is that we had very small groups. Probably would meet one half hour a day at least with those kids. And the groups would change, depending upon what focus it was. But usually, 90% of the time, it was those children that we thought were going to be partially proficient."

Example of school wide targeting

"At the beginning of the school year, the principal had the third grade CSAP results in reading and we targeted the kids at that point. We looked at their CSAP scores and figured that these were the kids that we could target and hopefully get up to the three (proficient). We did. We were aware. But we were also aware that you want to get your ones to twos because you also get points for that. And it is critical to get the threes to the fours because all that has impact on your overall score."

•Teachers reported targeting reading and writing at the partially proficient and unsatisfactory level.

"We looked at finding a collection of shorter passages and really focused on specific things with them. We really scrounged around the building trying to find quicker things that were really down and dirty that would illustrate certain themes rather than longer passes that didn't have that content. We wanted the students to be able to find the main idea, details, and make inferences from the material."

"The literacy resource teacher took some of the students who were partially proficient or unsatisfactory thirty minutes a day for pull-out."

While teachers often mentioned providing additional assistance to specific students in reading and writing, they rarely mentioned providing students with additional help in mathematics or science.

•Teachers reported targeting students with extra help and special programs to improve their CSAP scores.

Teachers used the following strategies to provide additional assistance to their targeted students: one-on-one, tutoring, resource teachers, and special programs. The number of special programs adopted by school districts was quite remarkable. For example, for reading and writing teachers mentioned being trained in the following programs: Six-Trait Writing, Lindamood-Bell, Read-to-Succeed, Step Up to Writing, Soar to Success, Roots and Wings, Read Naturally, and the William and Mary Language Arts Program. In math, Saxon Program, Sunshine Math, and Connected Mathematics, and the "Write" Way to Do Mathematics were discussed.

5th grade teacher
Advanced level: *We had math class for them.*
Proficient level: *(blank)*
Partially Proficient level: *We used Soar to Success with them.*
Unsatisfactory level: *(blank)*

3rd grade teacher
Advanced level: *Very little specific instructional effort*
Proficient level: *Extra instructional effort with content knowledge measured by the CSAP.*
Partially Proficient level: *This group was specifically targeted for additional instruction.*
Unsatisfactory level: *This group was not considered as a target for increased instruction with the content knowledge measured by the CSAP. This is not to say they were not helped to increase their reading skills.*

4th grade teacher
Advanced level: *(blank)*
Proficient level: *(blank)*
Partially Proficient level: *Small groups in reading to target summarizing*
Unsatisfactory level: *Worked with para pro one-on-one at least 2-3 times per week to target reading skills, i.e. retelling, summarizing, characterization.*

5th grade teacher
Advanced level: *enrichment program*
Proficient level: *(blank)*
Partially Proficient level: *HOTS program*
Unsatisfactory level: *Lindamood-Bell program*

Data from both the written and phone surveys suggest that the high-stakes testing environment in Colorado has influenced classroom instruction. Teachers are spending more time on test preparation activities and increasingly aligning their curriculum to match CSAP. Further, some teachers are targeted specific students with additional instruction to help improve overall school test scores. In the next section, teachers' perceptions of CSAP testing are discussed.

Teacher Perceptions

Teachers' perceptions of CSAP testing are important because as workers on the frontline, they must deal with the implementation and consequences of CSAP testing on a daily basis. To get at teachers' views about the CSAP, we asked them questions in two general areas: the usefulness of CSAP and support for instructional improvement. The response choices ranged from strong disagree to strongly agree. Teachers were also given the option to indicate not applicable if they believed the question did not apply to their teaching situation.

Table 12

Written survey question: Please indicate the extent to which you agree with each statement.

	Strongly disagree -2	 -1	Neutral 0	 1	Strongly Agree 2	N/A
CSAP results are helpful in identifying students' strengths and weakness.	31% SE=.028 Freq.=104	21% SE=.024 Freq.=75	15% SE=.022 Freq.=55	25% SE=.025 Freq.=89	5% SE=.012 Freq.=16	4% SE=.008 Freq.=13
CSAP is helping to improve schools.	38% SE=.020 Freq.=136	26% SE=.022 Freq.=92	17% SE=.025 Freq.=57	14% SE=.014 Freq.=53	3% SE=.007 Freq.=9	1% SE=.006 Freq.=5
My district and school provide me with the professional development I need to help all my students learn to high levels.	9% SE=.013 Freq.=30	14% SE=.016 Freq.=49	20% SE=.028 Freq.=75	34% SE=.022 Freq.=124	21% SE=.029 Freq.=71	1% SE=.006 Freq.=5
Low test scores on the CSAP help get additional resources to students with the greatest learning needs.	43% SE=.021 Freq.=159	24% SE=.020 Freq.=83	16% SE=.020 Freq.=53	13% SE=.014 Freq.=44	2% SE=.007 Freq.=7	2% SE=.006 Freq.=8

My school is more interested in increasing test scores than in improving overall student learning.	21% SE=.019 Freq.=69	24% SE=.020 Freq.=87	15% SE=.013 Freq.=53	20% SE=.017 Freq.=71	18% SE=.023 Freq.=70	1% SE=.006 Freq.=4
The items on the CSAP are a good representation of what students should know.	22% SE=.018 Freq.=73	24% SE=.027 Freq.=90	24% SE=.022 Freq.=82	23% SE=.028 Freq.=83	4% SE=.010 Freq.=15	3% SE=.010 Freq.=9
My district and school provide me with the resources (materials, time, etc.) I need to help all my students learn to high levels.	15% SE=.019 Freq.=50	19% SE=.014 Freq.=71	16% SE=.021 Freq.=53	32% SE=.026 Freq.=111	17% SE=.031 Freq.=63	1% SE=.006 Freq.=5
CSAP gives me important feedback about how well I am teaching in each curricular area.	49% SE=.029 Freq.=175	22% SE=.027 Freq.=74	14% SE=.017 Freq.=49	8% SE=.012 Freq.=27	1% SE=.004 Freq.=2	8% SE=.012 Freq.=27
CSAP helps to clarify which learning goals are most important.	40% SE=.027 Freq.=142	24% SE=.031 Freq.=84	18% SE=.017 Freq.=60	13% SE=.021 Freq.=50	1% SE=.004 Freq.=2	4% SE=.010 Freq.=15

As can be seen in table 12, items related to the utility of CSAP testing generally received disagreement ratings from teachers. More than half of the surveyed teachers strongly disagreed that CSAP results are helpful in identifying students' strengths and weaknesses, that CSAP is helping to improve schools, that the test gives important feedback to teachers about how well they are teaching in each curriculum area, or that it helps clarify important learning goals.

Teachers had a more positive view of their school and district's role in supporting them. Fifty-five percent of the teachers either agreed or strongly agreed that their districts were providing them with the professional development they needed to help all their students to high levels. Relatedly, 47% believed that their districts were providing them with the resources and materials they needed to help all their students learn.

Regarding school climate, 45% of the teachers either strongly disagreed or disagreed with the statement that their school was more interested in increasing test scores than in improving overall student learning. Only 38% either agreed or strongly agreed with this negative statement.

In the phone survey, teachers' responses followed the same pattern as the written survey with the exception of one question: in response to the statement that "CSAP results are helpful in identifying students' strengths and weaknesses," 46% of teachers

agreed or strongly agreed on the phone responses in contrast to 30% for the written survey (see Appendix I). Phone survey responses were consistent with the written survey for all the other categories. Quotations typifying teachers' perceptions of CSAP testing are listed below.

• **Teachers believe that CSAP has not helped to improve the quality of instruction.**

Quantitative data in Table 12 show that most teachers strongly disagree with the statements that "CSAP results are helpful in identifying students' strengths and weaknesses," "CSAP helps to clarify which learning goals are most important," and that "CSAP gives me important feedback about how well I am teaching in each curricular area." Most teachers also strongly disagree with the statement, "CSAP is helping to improve schools." In open-ended response items, teachers emphasized this point and elaborated upon it. Teachers overwhelmingly stated that CSAP has not helped to improve the quality of instruction for several reasons. First, CSAP content (i.e. test preparation and test taking strategies instruction or "teaching to the test") supplants other valuable content during the school year. Second, teachers and students, as a result of CSAP, are subjected to high amounts of stress and are demoralized. Third, as a result of CSAP, "fun" is eliminated from the curriculum and from the school day.

"For several weeks prior to CSAPs, we spent most of our language arts block practicing test booklets, released items, etc. We are focusing very narrowly on specific reading comprehension skills. For this extended period of time we are not reading good literature, working on grammar, contractions, possessive nouns, adjectives/adverbs, etc. We are basically focusing on skills needed to do well on the test. While these are valuable skills, we are having to short change many other valuable skills. We are not able to teach everything we have in the past. It is not a fair exchange."

"The stress put on staff and students to 'cram' for CSAP cost valuable instruction time resulting in poor instruction and increased disciplinary problems. Staff were irritable while testing was going on; students were frustrated and many skipped school with their parents approval based on media coverage of the CSAP."

"Focusing on CSAP preparation leaves less time for creative long term projects. The focus on CSAP scores creates an atmosphere of wary, dread -- not always a good atmosphere for learning."

"Negative influence comes through extreme pressure felt by teachers and handed down through administrators until "burnout" is happening so much earlier with teachers. We have narrowed our instruction to include large blocks of reading/writing time. Not a bad thing. The negative comes from excluding the "extras" which make education more fun -- and students learned without the drudgery. I'm glad I'm at the end of my career."

"CSAP places a large amount of stress on the students and teachers. This stress tends to decrease the focus and performance of the students."

BEST COPY AVAILABLE

• Teachers believe that CSAP has improved learning in some areas, at the expense of others.

Table 12 shows that very few teachers strongly agree with the statement "The items on the CSAP are a good representation of what students should know." Teachers went on to explain the results of this poor fit between what teachers believe should be taught and what is measured by the CSAP test. On several free response items teachers stated again and again that while reading, writing, and test-taking skills have improved, these gains have been at the expense of learning in the other standard content areas of science and social studies.

"I find that students coming up from elementary are grossly unprepared in areas of social studies and science. This is largely due to the myopic emphasis on areas assessed by CSAP at the elementary level."

"We were strongly discouraged to teach science and social studied for an entire quarter or so. After a 21/2-hour literacy block, we had to teach CSAP for an hour. Then all we had time for was math. In addition, higher order thinking skills were compromised for skill and drill. The prompts are vague and uninteresting. We taught to write for the CSAP versus real writing."

"Overload on language arts, less time for math, science, and social studies standards. the language arts overload has mostly focused on skills instruction and therefore decreased the time for genre studies, instruction with novels, and listening and speaking skills."

"A lot of time is spent on teaching students how to test. Time is used to teach students specific areas that will be tested and not on areas that are just as important for a well rounded student. Less time is spent on subjects that are not being tested like science - social studies."

• Teachers believe that the reporting of CSAP results (i.e. the School Report Card) is problematic.

As mentioned above, teachers believe that the items on the CSAP are not a good representation of what students should know. Interview and written response data reflecting teacher attitudes elaborate this point. Teachers often stated that they believe that the CSAP does not measure what is really going on in schools (i.e. does not offer a comprehensive vision of public education). Further, teachers stated that the school report card is an inappropriate, stressful, and demoralizing way to report on the performance of students, teachers, and schools.

"On the report cards, I think the grading thing is too reductionistic and it simplifies things a little too far. Schools are complex human organism and organizations of organism and soon as you assign a fairly simple mechanical or quantitative result, you reduce the complexity and by reducing the complexity, you minimize the potential of these humans who work together every day."

"The only thing that bothers me about CSAP in general is that it does not measure growth. So therefore kids that are lower may make extreme gains, but it doesn't reflect that. They don't get credit for that."

"There are so many other things that could be added to this assessment. The state is just using one test. There are so many other areas! Attendance records, scores on other types of tests, the number of students that go on to college, the advancement that kids are showing. You know, 25% advancement in students, raising themselves one grade level in a year is significant."

"How the results are used though is what I have a problem with. I think they are inappropriately used perhaps first by the media and then second by the state level government but mostly by the media. The way in which they are reported or used to make comparisons that are unhealthy for fostering positive community identity."

Data from the both the written and phone survey reflect the belief of many teachers that CSAP testing has not improved the overall quality of instruction in their classrooms. While teachers are focusing more carefully on reading and writing, they have less time to devote to science and social studies. In the next section, anticipated consequences of the school report card are discussed.

Consequences of School Report Card

In our pilot interviews with teachers, respondents predicted that the school report card would have negative effects on a variety of factors in the teaching environment. Therefore, we decided to ask teachers to predict what they thought the consequences of the public release of school report would be on such factors. For each factor, teachers were asked to predict if they thought it would decrease, stay the same, increase or to say that they had no opinion. Factors are related to: professional development, teachers leaving the profession, public regard for the teaching profession, faculty morale, and availability of resources to high and low performing schools (see Table 13).

Table 13

Written survey question: Based on your experience, what do you believe will be the consequences of the release of your school report card?

	Decrease	No change	Increase	No opinion
Availability of quality professional development to teach to standards will. . .	10% SE=.016 Freq.=35	54% SE=.032 Freq.=184	30% SE=.027 Freq.=112	6% SE=.006 Freq.=20
Availability of inservice training focused on test preparation will . . .	2% SE=.006 Freq.=5	26% SE=.030 Freq.=93	71% SE=.034 Freq.=249	2% SE=.006 Freq.=7

Number of poorly qualified teachers leaving the profession will . . .	5% SE=.011 Freq.=16	72% SE=.019 Freq.=257	17% SE=.013 Freq.=60	6% SE=.013 Freq.=21
Number of well qualified teachers leaving the profession will . . .	4% SE=.012 Freq.=13	19% SE=.019 Freq.=65	76% SE=.019 Freq.=269	1% SE=.006 Freq.=5
Public regard for the teaching profession will . . .	75% SE=.019 Freq.=266	19% SE=.018 Freq.=63	3% SE=.010 Freq.=12	3% SE=.007 Freq.=12
Faculty morale will . . .	88% SE=.022 Freq.=306	11% SE=.020 Freq.=38	1% SE=.004 Freq.=2	1% SE=.005 Freq.=5
Availability of resources to high performing schools will . . .	10% SE=.019 Freq.=36	26% SE=.025 Freq.=86	60% SE=.016 Freq.=212	5% SE=.009 Freq.=17
Availability of resources to low performing schools will . . .	50% SE=.020 Freq.=174	20% SE=.025 Freq.=76	25% SE=.017 Freq.=85	5% SE=.008 Freq.=17

The most striking finding to emerge from this section of the survey was the prediction by teachers that school report cards would have no effect on poorly qualified teachers but would cause well qualified teachers to leave the profession. Seventy-two percent of the teachers in the written survey thought that the school report card would have *no* impact on the number of *poorly qualified* teachers leaving the profession. In contrast, 76% of teachers believed that school report cards would increase the number of well qualified teachers leaving the profession. Similarly, 88% thought the release of school report cards would result in a decrease in faculty morale. 60% of the teachers in the written survey thought that there would be an increase in resource to high performing schools due to the CSAP while only 25% thought that there would be an increase in resources to low performing schools due to the report card. While only 30% of the teachers in the written survey believed that the availability of quality professional development to teach to standards would increase, 71% thought inservice training focused on test preparation would increase.

Quotations typifying teachers' beliefs about the release of the school report card are listed below.

• Teachers believe that well qualified teachers will leave the profession as a result of school report cards.

"A number of my friends are retiring because they don't want to be associated with a partially proficient school. And from their gut level, they are giving their all but getting a slap in the face. They know their kids are making advancements."

"I could say that a good half of the people in this building are talking about walking, about just giving it up and doing something where there is going to be more money and more prestige. We are getting tired of being spit on."

"I find that it is a demoralizing, stressful situation. Teachers are being judged year to year on completely different students. The pressure put on teachers has increased to the point where teachers will be leaving the profession."

"We can't even get teachers to come into schools like the one I am in because they are afraid that they will be called failing. Why should a young teacher come here when they can go to a wealthy school district? I mean the stigma that the grading has put on all schools in minority neighborhoods is just absolutely incredible. The talk that the good teachers don't come here, it basically frightens away anybody with any ability who doesn't know the community. Why should they come somewhere when they are going to be called a failure? I can't blame those teachers."

•Teachers believe that schools that perform poorly will have funds taken away.

"I think that the CSAP was a good idea to test the standards and to use it to find areas where we needed to improve and even where specific students needed to improve but I think tying it to the report cards has really made it a lot more political and put a lot of stress on teachers and students and on the test rather than on learning. I think just based on previous test results I think you can almost predict which schools and neighborhoods based on SES where they are going to have low scores and high scores. So I don't think it is really valid to say that then they are low and they should have their funds cut."

"Resources will be taken away from those schools, there's no doubt about it, its a stated plan, they're not going to be given extra resources, they're going to be shut down."

"It's a little ridiculous that schools that perform well are going to be rewarded, and schools that perform poorly are going to be punished."

•Teachers believe that public regard for the teaching profession will decline as result of the release of school report cards.

"I think school report cards are going to hurt public education. There is going to be a flight from public education into private education as a result of report cards. And the people who will remain in the public schools will be the people who can't afford the private schools. Consequently, the good teachers are also going to flee from public education to private schools."

BEST COPY AVAILABLE

"I think the test itself the way it was designed and what it was intended to do was good because it was intended to measure kids against standards. I think what happened is it became a political hot potato and it became a punitive weapon, and it is going to destroy us. Making us all charter schools is ludicrous. I don't see where they are going to go with that one at all. We don't have the parent involvement now. Making us a charter school isn't going to change that."

While teachers occasionally had positive things to say about the CSAP, their feelings toward the school report cards were uniformly negative. They felt it was an inappropriate way to guide schools and believed it would cause qualified teachers to leave the profession.

Conclusions

There is both good news and bad news in teachers' survey responses. The good news corroborates the rhetorical claims of the standards movement. Across Colorado teachers have responded to the content demands of the Colorado Model Content Standards by making substantial changes in curriculum and teaching practices. In almost all cases changes addressed to standards have served to make instruction more rigorous, and teachers' detailed descriptions suggest that these changes are more than superficial. The most striking finding has been the increased attention and effort devoted to writing across the state, which teachers attribute to both standards and CSAP testing. Six-Trait Writing was the most widely adopted new program, but in their efforts to improve both reading and writing instruction, schools districts adopted many new programs including Lindamood-Bell, Read-to-Succeed, Step Up to Writing, Soar to Success, Roots and Wings, Read Naturally, Power Writing, and the William and Mary Language Arts Program.

Teachers also attended to standards in mathematics instruction. They added units on probability and geometry as required by content standards and increased instructional activities involving problem solving and mathematical reasoning. There was also a dramatic increase in the number of teachers asking students to explain how they got their answers on math problems, because explanations would be required on CSAP. This is a positive example of teaching to the test, because being able to talk about and explain math problems is consistent with important learning goals. Many teachers also targeted instruction for specific groups of students, especially those identified as unsatisfactory or partially proficient. Teachers described numerous ways that they modified instruction especially to target reading and writing skills for partially proficient students. In some cases this increased attention occurred because standards made it clearer what instruction was needed, but it is also fair to say that the visibility of CSAP results may have heightened the motivation to attend to these needs.

The bad news reported by teachers is that CSAP also had negative effects on instruction that were not consistent with teaching to standards. Teachers considered the effects of CSAP to be harmful for two main reasons. First, increased time spent on reading, writing, and math reduced or eliminated time for science and social studies. Second, time spent preparing and practicing for CSAP was not a good use of instructional time even for the 3Rs. Practicing CSAP formats was not real reading or real math learning, and most importantly it was not real writing.

One teacher's response typifies teacher reports statewide about what they had to give up for CSAP:

> "Our district has told us to focus on reading, writing, and mathematics. Therefore, science and social studies, unless I can teach them in a reading, writing, or mathematics format, then they don't get taught. I don't teach science and social studies nearly as often and not purely as science and social studies. In the past I had hatched out baby chicks in the classroom as part of the science unit. I don't have time to do that. I have dissected body parts and I don't have time to do that. I have waited until now to start my economics unit for mini-society because it takes so much time. I can do it now because CSAP is over. We don't take as many field trips. We don't do community outreach like we used to, like visiting the nursing home or cleaning up the park (that we had adopted). Well, we don't have time for that any more."

Ironically, elimination of these activities works against the original intentions of standards-based reform. Not only are science and social studies significant areas of study in their own right, but the cognitive research that supported the development of standards in the first place showed that students would develop better conceptual understanding and be able to use their knowledge better if learning occurred in the context of real world applications.

Significant amounts of instructional time are spent preparing for CSAP tests. A minority of teachers statewide, between 20% and 30%, spent the few weeks before CSAP preparing their students by going over sample problems and administering practice tests. The majority of teachers, however, reported devoting much more time to test preparation activities saying that it went on, "regularly throughout the school year." Some of this practice was good subject matter instruction, like the math example above. But much of the test preparation was aimed at learning test formats and about what the test-makers wanted. And these things teachers did not countenance as good instruction.

The bad news also must include the negative effects of CSAP testing on faculty morale and teachers' anticipated fears about the School Report Card. Teachers do not believe that standards have negatively affected morale, but a great majority -- 81% of teachers -- believe that CSAP has hurt faculty morale. Teachers fear that the release of School Report Cards will further erode public confidence in education. They also believe that the consequences following from Report Cards will not affect poorly qualified teachers but will instead cause more well-qualified teachers to leave the profession. Some teachers self-reported that they might retire early because "teaching just isn't as much fun anymore."

The results of this survey pose a challenge and a paradox for policymakers trying to improve education: How to promote positive changes without so many negative side effects. This may well require more explicit consideration of the incentive systems and motivational aspects of educational reforms as well as attention to subject matter content. Some accountability proponents, like Frederick Hess cited earlier, insist that accountability systems must be "coercive," that educators must be rewarded or sanctioned on the basis of student performance or there won't be any change. Hess admits that this viewpoint runs counter to traditional values of the American educational system, which relied on the "good will of teachers and intrinsic motivation of students." Would policymakers be wise to follow this accountability advice?

While research on teachers' motivation is limited, there are extensive experimental studies on student motivation showing the negative consequences of relying on extrinsic rewards. Students who work for external rewards have less conceptual understanding, low confidence in themselves as learners, are unwilling to persist in trying to solve difficult problems, and stop working when rewards are removed. If this is what has been proven about student incentives, why would we want to adopt a similar system of rewards and punishments to motivate teachers?

By distinguishing between the positive and negative effects of standards-based reforms, teachers in our survey have pointed the way to more successful policies. More attention should be paid to content standards and to both professional development and curriculum materials that support instructional improvements. Less attention should be paid to raising test scores per se and to evaluating the quality of schooling only on the basis of CSAP results. Compared to traditional multiple-choice tests, CSAPs are pretty good tests. However, CSAPs are hardly adequate measures of how hard teachers are working or even of how much they are helping to improve student learning.

References

Bransford, J.D. Brown, A. L. & Cocking, R. R. (Eds.) (1999). *How people learn: brain, mind, experience, and school.* Washington, D.C.: National Academy Press.

Cannell, J. J. (1987). *Nationally normed elementary achievement testing in America's public schools. How all fifty states are above the national average.* Daniels, WV: Friends for Education.

Darling-Hammond and Wise, A. E. (1985). Beyond standardization: State standards and school improvement. *The Elementary School Journal, 85,* 315-336.

Elmore, R. F., & Rothman, R. (Eds.) (1999). *Testing, teaching, and learning: A guide for states and school districts.* Committee on Title I Testing and Assessment, Board on Testing and Assessment, National Research Council. Washington, DC: National Academy Press.

Hess, F. M. (2001, May). *Reform, resistance,...retreat? The predictable politics of accountability in Virginia.* Paper presented at the Brookings Conference on Accountability and Its Consequences for Students, Washington DC.

Kennedy, M. M., Ball, D. L. & McDiarmid, G. W. (1993). *A study package for examining and tracking changes in teachers' knowledge.* East Lansing, MI: National Center for Research on Teacher Education.

Koretz, D., Linn, R. L., Dunbar, S.B., & Shepard, L.A. (1991, April). *The effects of high-stakes testing on achievement: Preliminary findings about generalization across tests.* Paper presented at the annual meeting of the American Educational Research Association, Chicago.

Linn, R. L., Baker, E. L., & Dunbar, S. B. (1991). Complex, performance-based assessment: Expectations and validation criteria, *Educational Researcher, 20*(8), 15-21.

Linn, R. L., Graue, M. E., & Saunders, N.M. (1990). *Comparing state and district test results to national norms: Interpretations of scoring "Above the national average,"* CSE Technical Report 308. Los Angeles: Center for Research on Evaluation, Standards, and Student Testing.

Malcolm, S. M. (Ch.) (1993). *Promises to keep: Creating high standards for American students.* Washington, D.C.: National Education Goals Panel.

McLaughlin, M. W. & Shepard, L A. (1995). *Improving education through standards-based reform. A report of the National Academy of Educational Panel on Standards-Based Reform.* Stanford, CA: National Academy of Education.

National Commission on Excellence in Education. (1983). *A nation at risk: The imperative for educational reform.* Washington, D.C.: National Commission.

National Council on Education Standards and Testing. (1992). *Raising standards for American education: A report to Congress, the Secretary of Education, the National Education Goals Panel, and the American people.* Washington, D.C.: Author.

National Research Council. (1996). *National science education standards*. Washington D.C.: National Academy of Sciences.

O'Day, J. (in press). *Reconstitution as a remedy for school failure*. CPRE Policy Brief. Philadelphia: University of Pennsylvania, Consortium for Policy Research in Education.

Owens, B. (1999, December 8). "Putting Children First: A Plan for Safe and Excellent Public Schools." Denver, CO.

Pascoe, P. (2000, February 27). "Owens' reform plan is just a mirage." *The Denver Post*, K1, K6.

Popham, W. J. (1987). The merits of measurement-driven instruction. *Phi Delta Kappan, 68*, 697-682.

Porter, A. C. and Associates. (1994). *Reform of high school mathematics and science and opportunity to learn*. Washington D.C.: U.S. Department of Education, Office of Educational Research and Improvement, Educational Resources Information Center.

Resnick, L.B. & Resnick, D. P. (1992). Assessing the thinking curriculum: New tools for educational reform. In B. R. Gifford & M.C. O'Connor (Eds.), *Changing assessments: Alternative views of aptitude, achievement, and instruction*. Boston: Kluwer Academic Publishers.

Rottenberg, C. & Smith, M. L. (1990). *Unintended effects of external testing in elementary schools*. Paper presented at the annual meeting of the American Educational Research Association, Boston.

Shepard, L.A., & Dougherty, K. (1991, April). *Effects of high stakes testing on instruction*. Paper presented at the annual meeting of the American Educational Research Association, Chicago.

Smith, M. L. (1989). *The role of external testing in elementary schools*. Los Angeles: Center for Research on Evaluation, Standards, and Student Testing, University of California.

Smith, M.S. & O'Day, J (1990). Systemic school reform. *Politics of Education Association Yearbook 1990* (pp. 233-267). London: Taylor & Francis.

Stecher, B. M., Barron, S.L., Chun, T. and Ross, K. (2000). *The effects of the Washington state education reform on schools and classrooms*. CSE Technical Report 525. Los Angeles: UCLA Center for the Study of Evaluation.

Weisberg, H.F., Krosnick, J.A., & Bowen, B.D. (1996). *An introduction to survey research, polling, and data analysis, 3rd ed.* Thousand Oaks, CA: SAGE Publications.

Appendix A: Written survey

Teacher Survey on the CSAP and School Report Cards

This teacher survey is part of a national study of the effects of different state student testing and accountability systems on curriculum and instruction. Last year, the state legislature of Colorado passed legislation (SB 00-186) requiring that school report cards be issued for each school in the state. The purpose of this survey is to investigate and document effects of the school report card requirement and the extension of CSAP (Colorado Student Assessment Program) to grades 3 through 10. The research is funded by the U.S. Department of Education through the National Center for Evaluation, Standards and Student Testing.

Your participation in the study is completely voluntary. All responses are confidential and anonymous; DO NOT WRITE YOUR NAME ON THE QUESTIONNAIRE. Answers will be reported in summary form only (no specific individuals, schools, or districts will be identified from our data).

Your participation in this study is important. Results of this survey will be shared with the State Board of Education, Colorado Association of School Boards, Colorado Association of School Executives, Colorado Education Association and key policymakers in the state. The survey should take approximately 10 minutes to complete. If you should have any questions or concerns about the questionnaire, you may contact the researchers at (303) 492-5785. Confidential inquiries may be directed to the Executive Secretary, Human Research Committee, Graduate School, Campus Box 26, University of Colorado at Boulder, 80309. Thank you for your assistance.

Professor Robert Linn

University of Colorado at Boulder

P.S. If possible, please mail your completed questionnaire by April 10, 2001. Please also return the enclosed postcard. Your name on the separate postcard helps us keep track of who has responded to the survey while keeping the questionnaire anonymous. If you would like to receive a copy of the study results, please print your name and address on the bottom of the card. Thank you.

BACKGROUND INFORMATION

1. Including this year, how many years have you been teaching? _____

2. What is your gender? ____ female ____ male

3. Which of the following grade levels are you teaching this year? (Circle all that apply)

PK K 1 2 3 4 5 6 7 8 9 10 11 12

4. Which subjects do you teach? (Circle all that apply)

English/Language Arts Mathematics Social Studies Science

Other, please specify _____

5. Estimate your school's total enrollment _____

6. What is the name of your school district? _____

[This information will only be used to classify results by district's size and resource level]

7. The state will release school report cards in August based in large part on CSAP scores. What rating do you believe your school will be assigned? (Please check only one response).

___ excellent ___ high ___ moderate ___ low ___ unsatisfactory ___ you are unsure

8. Were any of your students assessed on the CSAP this year? Yes ____ No____*

[*If no, skip to question 10]

9. In which content area(s) were your students tested on the CSAP this year? _____

At which grade level(s)? _____

STANDARDS AND THE CSAP

10. To what extent have Standards-Based Reform and CSAP testing led to changes in the following areas? Please mark boxes in both the Standards and CSAP columns.

	With Standards			With CSAP Testing		
	Decreased	No Change	Increased	Decreased	No Change	Increased
Valuable professional development opportunities in the content areas have . . .						
Quality of classroom instruction has . . .						
Teaching higher-order thinking skills has . . .						
Faculty morale at my school has . . .						
Attention to lowest performing students has . . .						
Attention to students in the middle range has . . .						
Attention to highest performing students has . . .						
Student access to elective classes such as art, music, and physical education has . . .						
Amount of emphasis placed on writing has . . .						
Use of multidisciplinary approaches to subject matter has . . .						

INSTRUCTION

11. Describe your class. If you teach more than one class a day, describe the first class you teach each day where you have students who took the CSAP. **[If none of your students took the CSAP, describe the first class you teach each da**

Total number of students in class: _____
Estimate the number of students in this class who are limited or non-English proficient ____
Estimate the number of students in this class who participate in special education ____
Estimate the number of students in this class who qualify for free or reduced lunch_____

12. Compared with the rest of students at your school, are the majority of students in this class: (Check one)

____ in the top third of achievement for your subject area
____ in the middle third of achievement for your subject area
____ in the bottom third of achievement for your subject area

13. Based on your understanding of standards and proficiency levels, how many of the students in this class would you consider to be classified at each of the following proficiency levels in your subject area?

____ advanced ____ proficient ____ partially proficient ____ unsatisfactory

14. Thinking about the students you just listed in the four proficiency level categories, did you target any specific instructional efforts to help these students with the content knowledge measured by the CSAP? Yes___ No___

If yes, please describe:

Students at the advanced level: _____
_____

BEST COPY AVAILABLE

Students at the proficient level: _____

Students at the partially proficient level: _____

Students at the unsatisfactory level_____

15. Think about this class over the course of this school year. About how often do students in your class take part in the following activities? Please mark boxes in both the school year and CSAP columns.

	During the School Year			In Preparation for CSAP Test		
	Frequently (weekly)	Seldom (monthly)	Never	More often for CSAP	No Change for CSAP	Less Often for CSAP
Use rubrics to evaluate their work						
Discuss different ways that they solve particular problems						
Apply concepts or principles they have learned to real world situations						
Demonstrate their knowledge of basic skills or vocabulary						
Practice facts or procedures such as grammatical rules or formulas						
Read and discuss examples of good writing						
Conduct projects that extend over several days						
Conduct research in the library or on the internet						
Work in small groups on in-depth problem solving						
Use computers as part of regular instruction						
Write about how to solve a problem in an assignment or a test						

TEST PREPARATION ACTIVITIES

For items 16 to 21, try to recall how much time over this entire school year you spent in your classroom on the following test preparation activities. Then mark one box for each item.

16. Giving students review sheets of content you expected to be on the CSAP
no time a day or less 2-5 days 2-3 weeks 4 or more weeks

17. Giving students practice with the kinds of item formats that are on the CSAP
no time a day or less 2-5 days 2-3 weeks 4 or more weeks

18. Giving students commercially produced test preparation materials similar to the CSAP
no time a day or less 2-5 days 2-3 weeks 4 or more weeks

19. Giving students practice tests similar to the CSAP

no time a day or less 2-5 days 2-3 weeks 4 or more weeks

20. Instructing students on test-taking strategies

no time a day or less 2-5 days 2-3 weeks 4 or more weeks

21. When did most of the test preparation activities you conducted take place?

a few days before the CSAP Regularly throughout the school year

a few weeks before the CSAP Not applicable, I did not do any test preparation.

TEACHER PERCEPTIONS

22. Please indicate the extent to which you agree with each statement. Circle one number in each row.

	Strongly Disagree				Strongly Agree	Not Applicable
CSAP results are helpful in identifying students strengths and weakness.	-2	-1	0	+1	+2	N/A
CSAP is helping to improve schools.	-2	-1	0	+1	+2	N/A
My district and school provide me with the professional development I need to help all my students learn to high levels.	-2	-1	0	+1	+2	N/A
Low test scores on the CSAP help get additional resources to students with the greatest learning needs.	-2	-1	0	+1	+2	N/A
My school is more interested in increasing test scores than in improving overall student learning.	-2	-1	0	+1	+2	N/A
The items on the CSAP are a good representation of what students should know.	-2	-1	0	+1	+2	N/A
My district and school provide me with the resources (materials, time, etc.) I need to help all my students learn to high levels.	-2	-1	0	+1	+2	N/A
CSAP gives me important feedback about how well I am teaching in each curricular area.	-2	-1	0	+1	+2	N/A
CSAP helps to clarify which learning goals are most important.	-2	-1	0	+1	+2	N/A

CONSEQUENCES OF SCHOOL REPORT CARD

23. Based on your experience, what do you believe will be the consequences of the release of your school's report card? Please mark one box in each row.

	Decrease	No Change	Increase	No Opinion
Availability of quality professional development to teach to standards will. .				
Availability of inservice training focused on test preparation will . . .				
Number of poorly qualified teachers leaving the profession will . . .				
Number of well qualified teachers leaving the profession will . . .				
Public regard for the teaching profession will . . .				
Faculty morale will . . .				
Availability of resources to high performing schools will . . .				
Availability of resources to low performing schools will . . .				

24. If you teach students who took the CSAP, give one or two examples of how CSAP has helped you to improve the quality of instruction in your classroom.

_____

_____

_____

BEST COPY AVAILABLE

_____

_____

25. If you teach students who took the CSAP, give one or two examples of how CSAP has had a negative influence on the quality of instruction in your classroom.

_____

_____

_____

_____

_____

Thank you for taking the time to fill out this survey. Please return it in the envelope provided.

Appendix B: Telephone survey

Teacher ID _____
District Name _____
Elementary/ Middle/ H.S. _____

Before we begin, I would like to ask permission to tape-record your answers to open-ended questions. This will facilitate my note taking. After I complete the interview, I will transcribe your responses and then erase the tape. Would that be ok?

Yes ___ *No* ___

Are you ready to begin? First, I have a few background questions to ask you.

1. Including this year, how many years have you been teaching? _____

2. **(Interviewer mark gender: 1= female, 2 = male)** _____

3. What grade levels are you teaching this year? **(Circle all that apply)**
2 3 4 5 6 7 8 9 10 11 12 _____

4. What subjects do you teach? **(Circle all that apply)**
English/Language Arts (1) Mathematics (2) Social Studies (2) Science (4)
Elementary (5 – all subjects)
Other, please specify _____ _____

5. Please estimate your school's total enrollment _____

7. In August, the state will release school report cards based in large part on CSAP scores. Which one of the following ratings do you believe your school will be assigned?

excellent (5) high (4) moderate (3) low (2) unsatisfactory (1) you are unsure (0) _____

8. a. Were any of your students assessed on the CSAP this year? Yes ____ No___
b. If yes, in which content areas were your students assessed on the CSAP? _____
c. at which grade levels? _____

9. Do you teach more than one class a day? Yes (1) No (2) [*Teachers that stay with the same students all day have one class*] _____

10. What is the total number of students in your class? _____
Can you estimate the number of students in this class who are limited or non-English proficient? _____
Can you estimate the number of students in this class who participate in special education? _____
Can you estimate the number of students in this class who qualify for free and reduced lunch? _____

11. Compared with the rest of students <u>at your school</u>, are the majority of students in this class: _____
(3) in the top third of achievement for your subject area
(2) in the middle third of achievement for your subject area
(1) in the bottom third of achievement for your subject area

12. The next set of questions are about standards-based reforms.

Have you made any changes in your classroom instruction in the last few years in response to Standards-Based Reforms?

(1)Yes (ask #13)
(0)No (skip to #14) ____

13. Can you describe one or two key changes?

14.-20. Ok, thinking about your class over the course of this school year, how often do your students take part in the following activities? Say whether each activity is done frequently (meaning weekly), seldom (monthly) or never.

	During the School Year		
	Frequently (weekly)	Seldom (monthly)	Never
14. The first activity is having students use rubrics to evaluate their work. Do students do this frequently, seldom or never?			
15. **How often do students** . . . apply concepts or principles they have learned to real world situations?			
16. demonstrate their knowledge of basic skills or vocabulary			
17. practice facts or procedures such as grammatical rules or formulas			
18. conduct projects that extend over several days			
19. work in small groups on in-depth problem solving			
20. write about how to solve a problem			

21. This next question focuses on CSAP testing. Has CSAP testing changed **the content** of instruction in your classroom?

(1)Yes (ask #22)
(0)No (skip to #23) ____

22. Can you give one or two examples of how CSAP has changed instruction in your classroom?

23-30. Next I have some questions about changes related to both Standards-Based Reform and CSAP testing. I am going to ask whether each of the following items has increased, decreased, or stayed the same in response to standards. Then I will go back over the list again and ask if the items have changed as a result of CSAP testing.

	With Standards			**With CSAP Testing**		
	Increased	No Change	Decreased	Increased	No Change	Decreased
23. The first item is valuable professional development opportunities in the content areas. **Would you say it has increased, decreased, or stayed the same as a result of standards?**						
24. How about the quality of classroom instruction ?						
25. teaching higher-order thinking skills						

26. attention to the lowest performing students		
27. attention to the highest performing students		
28. student access to elective classes such as art, music, and physical education		
29. amount of emphasis placed on writing		
30. use of multidisciplinary approaches to subject matter		

Now we are going to go back through that list and I want you to think about these items in terms of changes in response to CSAP testing. Would you say that in response to CSAP testing, valuable professional development opportunities in the content areas have increased, decreased or stayed the same?

The next set of questions is about test preparation.
For each example of a test preparation activity, I'm going to ask if you spent no time on it, a day or less, 2-5 days, 2-3 weeks, or 4-8 weeks.

31. The first example is: giving students review sheets of content you expected to be on the CSAP. Did you spend:
(0) no time (1) a day or less (2) 2-5 days (3) 2-3 weeks (4) 4 -8 weeks ____

32. Giving students practice with the kinds of item formats that are on the CSAP
(0) no time (1) a day or less (2) 2-5 days (3) 2-3 weeks (4) 4 -8 weeks ____

33. Giving students commercially produced test preparation materials similar to the CSAP
(0) no time (1) a day or less (2) 2-5 days (3) 2-3 weeks (4) 4 -8 weeks ____

34. Giving students practice tests similar to the CSAP
(0) no time (1) a day or less (2) 2-5 days (3) 2-3 weeks (4) 4 -8 weeks ____

35. Instructing students on test-taking strategies
(0) no time (1) a day or less (2) 2-5 days (3) 2-3 weeks (4) 4 -8 weeks ____

36. When did most of the test preparation activities you conducted take place?
(1) a few days before the CSAP (3) Regularly throughout the school year
(2) a few weeks before the CSAP (4) Not applicable, I did not do any test preparation ____

37. Did you eliminate any activities from your schedule to allow time for test preparation activities?
(1) Yes (0) No ____

(If yes) what activities did you eliminate?

38. The next two questions focus on content standards and proficiency levels.
Based on your understanding of standards and proficiency levels, how many of the students in your class would you consider to be classified at each of the following proficiency levels in your subject area?

____ advanced ____ proficient ____ partially proficient ____unsatisfactory

39. Thinking about the students you just described in the four proficiency level categories, did you target any specific instructional efforts to help students with the *content knowledge* measured by the CSAP?

(1)Yes (ask #40) ____
(0)No (skip to #41)

40a. For example, please describe any specific instructional efforts for students at the advanced level?

b. How about students at the proficient level?

c. Students at the partially proficient level:

d. And finally, how about students at the unsatisfactory level?

e. (If teacher says she/he focused instructional efforts on the whole class then record information under 40e)

41. We are almost finished. The next set of questions focus on the uses and effects of CSAP. There are six response choices ranging from strongly agree, agree, neutral, to disagree and strongly disagree. You may also say not applicable if the situation does not apply to your teaching situation.

	Strongly Agree				Strongly Disagree	Not Applicable
The first statement says, "CSAP results are helpful in identifying students' strengths and weaknesses." The response choices are strongly agree, agree, neutral, disagree, or strongly disagree.	+2	+1	0	-1	-2	N/A
CSAP is helping to improve schools.	+2	+1	0	-1	-2	N/A
My district and school provide me with the professional development I need to help all my students learn to high levels..	+2	+1	0	-1	-2	N/A
Low test scores on the CSAP help get additional resources to students with the greatest learning needs. The response choices are strongly agree, agree, neutral, disagree, or strongly disagree.	+2	+1	0	-1	-2	N/A
My school is more interested in increasing test scores than in improving overall student learning.	+2	+1	0	-1	-2	N/A
The items on the CSAP are a good representation of what students should know.	+2	+1	0	-1	-2	N/A
My district and school provide me with the resources I need to help all my students learn to high levels (for example: materials)	+2	+1	0	-1	-2	N/A

42. This is the last set of questions. They focus on school report cards. Based on your experience, what do you believe will be the consequences of the release of your school's report card on each of the following? Will they increase, decrease, stay the same, or you have no opinion?

	Increase	No Change	Decrease	No Opinion
The first item is Availability of quality professional development to teach to standards Do you believe it will increase, stay the same, decrease or you have no opinion?				
Availability of inservice training focused on test preparation				
Number of poorly qualified teachers leaving the profession				
Number of well qualified teachers leaving the profession				
Public regard for the teaching profession				
Faculty morale				
Availability of resources to high performing schools				
Availability of resources to low performing schools				

43. Are there any other comments that you would like to make about CSAP or the school report card?

Thank you for taking the time to complete this survey. If you would like a copy of results, please provide me with your address and we will send you a copy.

Appendix C: Qualitative data coding

Codes for Written Survey #14

Code Name

iec — Instructional Efforts: Content Related

- w — writing
- r — reading
- m — math
- s — science

- iec-h — higher level thinking/project-oriented/more challenging
- iec-l — basic skills/remediation/drill type activities

- tts — test taking skills/strategies
  - Reading directions
  - Answer development (i.e., complete sentences, organized response, answering all parts of questions)
  - Practice Test/sample questions
  - Vocabulary
  - Guide Books
  - Writing to prompt
  - Formatting (e.g., response types, short answer)
  - Rubric
  - More frequent assessment

- eh — Extra Help
  - one/one
  - tutoring
  - resource teacher, para-professionals, SPED, ESL, Title I
  - gifted and talented (specialist-high)
  - ability grouping
  - small groups
  - slowed down instruction
  - Special Program (e.g. 6-Trait, Step up to Writing)

tag — Targeting a group

- a — advanced
- p — proficient
- pp — partially proficient
- u — unsatisfactory
- ig — Ignore a group

Codes for #13 on Phone Survey

Code Name

acs	aligning curriculum/instruction/lessons with standards
acsadd	something was added
acselim	something was eliminated
ass	more and different types of assessment
acc	aligning curriculum with CSAP
accadd	something was added
accelim	something was eliminated
commtch	staff is involved in discourse regarding standards
commpar	parents are involved in discourse regarding standards
commstud	students are involved in discourse regarding standards
efun	eliminating favorite/fun
mwr	more emphasis on writing
awarestud	students have a heightened awareness of standards
awaretch	teachers have a heightened awareness of standards

Codes for #22 on Phone Survey

Code Name

acs		aligning curriculum/instruction/lessons with standards
	acsadd	something was added
	acselim	something was eliminated
ass		more and different types of assessment
acc		aligning curriculum with CSAP
	accadd	something was added
	-R	Reading content added
	-W	Writing content added
	-M	Math content added
	-S	Science content added
	-TPRAC	CSAP-similar practice materials added (format, guidebooks, modified quizzes, etc.)
	-TSTRAT	Test strategies instructed
	-PROC	Process emphasized
	-PROG	Special program or instructional materials (other than TPRAC) added
	-GRP	Ability groups / small groups added
	accelim	something was eliminated
	-SCI	Science content eliminated
	-SOC	Social studies content eliminated
	-LAB	Labs eliminated
	-TIM	Instructional time decreased
	-ACT	Activities eliminated
	-GAM	Games eliminated
	-FIELD	Field trip eliminated
	-THEM	Thematic unit eliminated
	-HON	Hands-on activities eliminated
	NAR	"Narrowing of the curriculum"
	-DEPTH	Cover more material with less detail
	-PRJ	Project eliminated
	EFUN	eliminating favorite/fun

commtch	staff is involved in discourse regarding standards
commpar	parents are involved in discourse regarding standards
commstud	students are involved in discourse regarding standards
awarestud	students have a heightened awareness of standards
awaretch	teachers have a heightened awareness of standards

Codes for #37 on Phone Survey

Code Name

acc

accelim something was eliminated

-R	Reading content eliminated
-INT	Integrated/Multidisciplinary/Thematic content or units eliminated
-ELECT	Elective content eliminated (Art, Drama, Music, PE, Library)
-TECH	Technology content eliminated / less time devoted
-COMMSRV	Community Service projects eliminated
-ORAL	Oral presentations eliminated
-W	Writing content eliminated
-M	Math content eliminated
-S	Science content eliminated
-SOC	Social studies content eliminated
-LAB	Labs eliminated
-TIM	Instructional time decreased
-ACT	Activities eliminated
-GAM	Games eliminated
-FIELD	Field trip eliminated
-HON	Hands-on activities eliminated
-DEPTH	Cover more material with less detail
-PRJ	Project eliminated
EFUN	eliminating favorite/fun

Appendix D: Qualitative data theme counts

Question 13 on Phone Survey: Changes made in classroom instruction in the last few years in response to standards-Based Reforms

Total responses: 136	Curriculum was aligned with standards	More and different types of assessment were implemented	Curriculum was aligned with CSAP	Discourse regarding Standards was opened	Fun was eliminated	There was a greater emphasis on writing	Students and teachers had a heightened awareness of Standards
Number of responses in each theme	101	26	23	22	3	12	9

Question 14 on Written Survey: Instructional efforts targeted to help students at each of the four proficiency levels on the content knowledge measured by the CSAP

Total responses: 76	Instructional efforts toward writing content	Instructional efforts toward reading content	Instructional efforts toward math content	Instructional efforts toward higher level thinking content	Instructional efforts toward basic skills and remediation content	Extra help (one-on-one, small groups, etc)
Number of responses in each theme for a: advanced (total: 37)	9	12	4	24	1	13
Number of responses in each theme for b: proficient (total: 25)	9	13	2	7	0	6
Number of responses in each theme for c: partially-proficient (total: 54)	14	26	2	2	4	33
Number of responses in each theme for d: unsatisfactory (total: 52)	13	21	2	0	4	41

Question 22 on Phone Survey: How CSAP has changed instruction in teacher's classroom

Total responses: 100	Curriculum was added in alignment with CSAP	Curriculum was eliminated in alignment with CSAP	Curriculum, instruction, or lessons were aligned with Standards	More and different types of assessment were added	Discourse regarding Standards was opened	Students and teachers had a heightened awareness of Standards
Number of responses in each theme	86	82	7	8	3	2

Question 24 on Written Survey: Examples of how CSAP helped the teacher to improve the quality of instruction in his or her classroom

Total responses: 261	There was no improvement or quality stayed the same	There were improvements related to content	There were improvements to the structure of teaching, classroom, or school	There were improvements in test taking strategies instruction
Number of responses in each theme	50	87	63	32

Question 25 on Written Survey: Examples of how CSAP had a negative influence on the quality of instruction in teacher's classroom

Total responses: 285	CSAP (in general) has taken time away from instruction	Content was shortened or eliminated	Too much time was spent on test preparation and test taking skills instruction	Quality of instruction suffered because of CSAP	Creativity lost, fun, projects, or field trips eliminated	Stress, anxiety, pressure or fear increased due to CSAP	Morale decreased as a result of CSAP
Number of responses in each theme	69	38	44	33	43	78	48

Question 37 on Phone Survey: Activities eliminated from teacher's schedule to allow time for test preparation activities

Total responses: 91	Material eliminated in the content areas (e.g. math, science, etc.)	Instructional formats eliminated (e.g. games, field trips, projects, etc.)	Fun, depth, or instructional time decreased or eliminated
Number of responses in each theme	67	32	22

Question 40 on Phone Survey: Instructional efforts targeted to help students at each of the four proficiency levels on the content knowledge measured by the CSAP

Total responses: 57	Instructional efforts related to content	Instructional efforts related to test taking skills and strategies	Assignments and assessments were aligned with CSAP	Extra help (e.g. tutoring, pull out, or resource teachers) was provided	One of the four proficiency levels was explicitly targeted
Number of responses in each theme for a: advanced (total: 18)	11	1	8	4	1
Number of responses in each theme for b: proficient (total: 18)	6	9	2	9	2
Number of responses in each theme for c: partially-proficient (total: 41)	12	9	3	29	8
Number of responses in each theme for d: unsatisfactory (total: 26)	10	5	1	17	3

Question 43 on Phone Survey: Additional comments about CSAP or the school report card

Total responses: 107	CSAP results are being used inappropriately	School Report Card is an inappropriate way to measure students, teachers, and schools	CSAP is unfair to certain populations	CSAP is good	Account ability is good	CSAP does not give a comprehensive view of students, teachers, and schools	CSAP test takes too long
Number of responses in each theme	10	28	16	12	5	30	8

CSAP preparation takes too long	CSAP/School Report Card is stressful and demoralizing to teachers and students	There is no accountability for students with CSAP	CSAP helps teachers to focus on the standards	CSAP is narrowing the curriculum	CSAP test is good, but the School Report Card is Bad	Discrepancy in resources causes an unequal playing field
5	26	11	6	7	19	18

Appendix E: Cross tabulated written survey data for CSAP testing versus standards-based reform

Written survey data by self-reported school report card grade:

Due to standards-based reform and CSAP testing, faculty morale at my school has...
1=Decreased
2=No change
3=Increased

School Report Card Grades						
	Standards			CSAP		
	N	Mean	SD	N	Mean	SD
Excellent	26	1.584	0.100	26	1.248	0.089
High	103	1.774	0.084	100	1.213	0.063
Moderate	116	1.758	0.071	115	1.264	0.045
Low	63	1.755	0.071	62	1.133	0.032
Unsatisfactory	25	1.816	0.155	25	1.080	0.001

Due to standards-based reform and CSAP testing, the amount of emphasis placed on writing has...
1=Decreased
2=No change
3=Increased

School Report Card Grades						
	Standards			CSAP		
	N	Mean	SD	N	Mean	SD
Excellent	26	2.661	0.122	26	2.458	0.173
High	103	2.761	0.045	101	2.788	0.077
Moderate	116	2.753	0.032	113	2.800	0.032
Low	63	2.734	0.055	61	2.777	0.063
Unsatisfactory	25	2.773	0.045	25	2.607	0.063

Due to school report cards, faculty morale will...
1=Decreased
2=No change
3=Increased

School Report Card Grades			
	N	Mean	SD
Excellent	27	1.309	0.084
High	106	1.236	0.063
Moderate	121	1.143	0.001
Low	62	1.079	0.032
Unsatisfactory	25	1.000	0.001

Written survey data by school type:

Due to standards-based reform and CSAP testing, faculty morale at my school has...
1=Decreased
2=No change
3=Increased

School Type						
	Standards			CSAP		
	N	Mean	SD	N	Mean	SD
Elementary	164	1.773	0.063	163	1.189	0.032
Middle School	78	1.714	0.045	76	1.165	0.055
High School	80	1.797	0.063	78	1.233	0.045
Mix	21	1.762	0.072	19	1.349	0.037

Due to standards-based reform and CSAP testing, the amount of emphasis placed on writing has...
1=Decreased
2=No change
3=Increased

School Type						
	Standards			CSAP		
	N	Mean	SD	N	Mean	SD
Elementary	164	2.763	0.045	162	2.756	0.071
Middle School	78	2.758	0.045	77	2.782	0.045
High School	80	2.720	0.055	77	2.757	0.055
Mix	20	2.747	0.056	22	2.644	0.095

Due to standards-based reform and CSAP testing, faculty morale will...
1=Decreased
2=No change
3=Increased

School Type			
	N	Mean	SD
Elementary	168	1.149	0.045
Middle School	79	1.113	0.032
High School	82	1.210	0.055
Mix	22	1.197	0.090

Written survey data by teacher type:

Due to standards-based reform and CSAP testing, faculty morale at my school has...
1=Decreased
2=No change
3=Increased

Teacher Type						
	Standards			CSAP		
	N	Mean	SD	N	Mean	SD
Elective	40	1.777	0.114	40	1.220	0.063
Specialist	31	1.725	0.095	30	1.192	0.045
Math	59	1.823	0.095	58	1.134	0.032
Science	49	1.812	0.071	48	1.157	0.055
Social Studies	45	1.756	0.089	42	1.190	0.055
English/LA	85	1.772	0.045	83	1.292	0.055

Due to standards-based reform and CSAP testing, the amount of emphasis placed on writing has...
1=Decreased
2=No change
3=Increased

Teacher Type						
	Standards			CSAP		
	N	Mean	SD	N	Mean	SD
Elective	40	2.762	0.063	40	2.828	0.071
Specialist	31	2.854	0.045	31	2.913	0.001
Math	59	2.784	0.045	57	2.859	0.045
Science	49	2.703	0.063	48	2.760	0.071
Social Studies	45	2.748	0.055	43	2.791	0.063
English/LA	85	2.708	0.055	84	2.729	0.045

Due to standards-based reform and CSAP testing, faculty morale will...
1=Decreased
2=No change
3=Increased

School Type			
	N	Mean	SD
Elective	41	1.126	0.032
Specialist	32	1.120	0.077
Math	61	1.045	0.032
Science	49	1.086	0.001
Social Studies	47	1.267	0.084
English/LA	87	1.134	0.045

Appendix F: Telephone survey data for CSAP testing versus standards-based reform

	With Standards			With CSAP Testing		
	Decreased (0) Recoded-1	No change (1) Recoded-2	Increased (2) Recoded-3	Decreased (0) Recoded-1	No change (1) Recoded-2	Increased (2) Recoded-3
The first item is valuable professional development opportunities in the content areas. Would you say it has increased, decreased, or stayed the same as a result of standards?	5.4% SE=.015 Freq.=8	41.2% SE=.050 Freq.=60	53.4% SE=.052 Freq.=83	17.3% SE=.029 Freq.=26	38.4% SE=.032 Freq.=59	44.2% SE=.041 Freq.=69
How about the quality of classroom instruction ?	4.7% SE=.018 Freq.=8	44.6% SE=.053 Freq.=64	50.6% SE=.052 Freq.=84	20.2% SE=.035 Freq.=30	49.8% SE=.055 Freq.=73	30.0% SE=.038 Freq.=51
Teaching higher-order thinking skills	2.2% SE=.015 Freq.=3	53.9% SE=.035 Freq.=80	43.9% SE=.039 Freq.=72	10.9% SE=.028 Freq.=16	48.5% SE=.042 Freq.=72	40.6% SE=.050 Freq.=69
Attention to the lowest performing students	6.4% SE=.015 Freq.=11	44.5% SE=.033 Freq.=68	49.1% SE=.041 Freq.=78	10.3% SE=.025 Freq.=15	42.0% SE=.049 Freq.=64	47.7% SE=.047 Freq.=77
Attention to the highest performing students	11.8% SE=.030 Freq.=17	63.2% SE=.032 Freq.=95	24.9% SE=.034 Freq.=43	15.3% SE=.032 Freq.=23	60.6% SE=.034 Freq.=93	24.1% SE=.034 Freq.=41
Student access to elective classes such as art, music, and physical education	16.2% SE=.008 Freq.=24	74.9% SE=.049 Freq.=114	8.9% SE=.035 Freq.=14	21.8% SE=.032 Freq.=34	75.9% SE=.033 Freq.=115	2.3% SE=.009 Freq.=4
Amount of emphasis placed on writing	1.6% SE=.152 Freq.=3	15.2% SE=.028 Freq.=22	83.2% SE=.029 Freq.=130	1.9% SE=.011 Freq.=3	16.1% SE=.029 Freq.=24	81.9% SE=.030 Freq.=128
Use of multidisciplinary approaches to subject matter	10.2% SE=.021 Freq.=14	52.4% SE=.029 Freq.=77	37.4% SE=.035 Freq.=63	18.7% SE=.028 Freq.=27	47.5% SE=.044 Freq.=69	33.8% SE=.047 Freq.=58

Appendix G: Written survey data for instruction during the school year

School Year Instructional Practices

Survey Question: ***How often do students in your class take part in the following activities during the school year?***

	During the School Year		
	Frequently	Seldom	Never
Use rubrics to evaluate their work	40% SE=.026 Freq.=136	48% SE=.023 Freq.=171	12% SE=.020 Freq.=38
Discuss different ways that they solve particular problems	82% SE=.017 Freq.=285	16% SE=.015 Freq.=54	2% SE=.005 Freq.=8
Apply concepts or principles they have learned to real world situations	84% SE=.017 Freq.=288	16% SE=.017 Freq.=59	0% SE=.003 Freq.=2
Demonstrate their knowledge of basic skills or vocabulary	91% SE=.012 Freq.=316	9% SE=.012 Freq.=33	0% SE=0 Freq.=0
Practice facts or procedures such as grammatical rules or formulas	69% SE=.023 Freq.=236	25% SE=.023 Freq.=90	6% SE=.009 Freq.=20
Read and discuss examples of good writing	63% SE=.019 Freq.=223	29% SE=.018 Freq.=98	8% SE=.008 Freq.=28
Conduct projects that extend over several days	57% SE=.025 Freq.=191	40% SE=.026 Freq.=146	3% SE=.013 Freq.=12
Conduct research in the library or on the internet	32% SE=.029 Freq.=103	55% SE=.030 Freq.=196	13% SE=.015 Freq.=47
Work in small groups on in-depth problem solving	54% SE=.036 Freq.=191	40% SE=.040 Freq.=137	6% SE=.011 Freq.=22
Use computers as part of regular instruction	48% SE=.021 Freq.=161	39% SE=.025 Freq.=141	12% SE=.025 Freq.=47
Write about how to solve a problem in an assignment or a test	48% SE=.024 Freq.=162	42% SE=.028 Freq.=151	10% SE=.016 Freq.=34

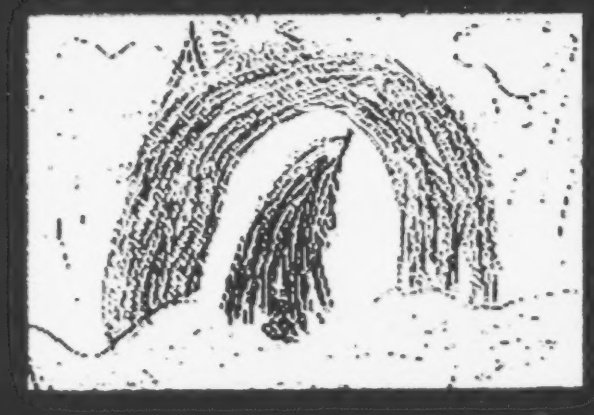
Appendix H: Written survey data for instructional changes due to CSAP test preparation

Instructional Practices in Preparation for CSAP Testing
Survey question: ***How often do students in your class take part in the following activities in preparation for CSAP testing?***

	In Preparation for CSAP Testing		
	More Often	No Change	Less Often
Use rubrics to evaluate their work	17% SE=.028 Freq.=51	75% SE=.036 Freq.=251	8% SE=.014 Freq.=26
Discuss different ways that they solve particular problems	24% SE=.022 Freq.=79	71% SE=.019 Freq.=234	5% SE=.010 Freq.=15
Apply concepts or principles they have learned to real world situations	9% SE=.016 Freq.=31	79% SE=.034 Freq.=257	12% SE=.025 Freq.=39
Demonstrate their knowledge of basic skills or vocabulary	31% SE=.018 Freq.=106	66% SE=.021 Freq.=213	3% SE=.012 Freq.=9
Practice facts or procedures such as grammatical rules or formulas	30% SE=.029 Freq.=103	64% SE=.028 Freq.=206	6% SE=.009 Freq.=18
Read and discuss examples of good writing	36.5% SE=.024 Freq.=116	58% SE=.021 Freq.=194	5.5% SE=.010 Freq.=17
Conduct projects that extend over several days	8% SE=.012 Freq.=26	70% SE=.032 Freq.=229	22% SE=.033 Freq.=73
Conduct research in the library or on the internet	4% SE=.010 Freq.=13	74% SE=.033 Freq.=240	22% SE=.029 Freq.=72
Work in small groups on in-depth problem solving	9% SE=.016 Freq.=27	72% SE=.034 Freq.=236	19% SE=.031 Freq.=63
Use computers as part of regular instruction	5% SE=.011 Freq.=16	77% SE=.028 Freq.=252	18% SE=.026 Freq.=58
Write about how to solve a problem in an assignment or a test	26% SE=.025 Freq.=92	68% SE=.024 Freq.=219	6% SE=.013 Freq.=18

Appendix I: Telephone survey data for teacher perceptions of CSAP testing

	Strongly Disagree				Strongly Agree	
	-2	-1	0	+1	+2	N/A
The first statement says, "CSAP results are helpful in identifying students' strengths and weaknesses." The response choices are strongly agree, agree, neutral, disagree, or strongly disagree.	15.7% SE=.028 Freq.=26	15.9% SE=.031 Freq.=24	19.5% SE=.041 Freq.=29	34.1% SE=.038 Freq.=60	5.8% SE=.016 Freq.=10	9.0% SE=.023 Freq.=12
CSAP is helping to improve schools.	16.5% SE=.031 Freq.=26	30.8% SE=.046 Freq.=48	23.9% SE=.027 Freq.=38	23.2% SE=.031 Freq.=39	3.3% SE=.013 Freq.=6	2.4% SE=.010 Freq.=4
My district and school provide me with the professional development I need to help all my students learn to high levels..	5.5% SE=.023 Freq.=7	14.9% SE=.025 Freq.=24	15.6% SE=.032 Freq.=23	43.2% SE=.053 Freq.=72	20.8% SE=.036 Freq.=35	0% Freq.=0
Low test scores on the CSAP help get additional resources to students with the greatest learning needs. The response choices are strongly agree, agree, neutral, disagree, or strongly disagree.	21.0% SE=.025 Freq.=33	34.8% SE=.043 Freq.=61	8.0% SE=.026 Freq.=13	14.7% SE=.031 Freq.=25	1.9% SE=.010 Freq.=3	19.6% SE=.058 Freq.=26
My school is more interested in increasing test scores than in improving overall student learning.	24.8% SE=.040 Freq.=39	37.6% SE=.050 Freq.=62	7.0% SE=.025 Freq.=10	22.8% SE=.030 Freq.=37	5.5% SE=.016 Freq.=9	2.3% SE=.011 Freq.=4
The items on the CSAP are a good representation of what students should know.	12.3% SE=.022 Freq.=20	30.9% SE=.063 Freq.=48	14.6% SE=.027 Freq.=23	30.6% SE=.040 Freq.=50	4.9% SE=.013 Freq.=8	6.7% SE=.020 Freq.=11
My district and school provide me with the resources I need to help all my students learn to high levels (for example: materials)	7.2% SE=.023 Freq.=9	19.0% SE=.029 Freq.=31	14.1% SE=.032 Freq.=20	45.6% SE=.047 Freq.=78	13.5% SE=.019 Freq.=22	0.6% SE=.007 Freq.=1



U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

NOTICE

Reproduction Basis

X	This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.
	This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").